

# Estado nutricional y adherencia a la dieta mediterránea en población mayor de 40 años. Utilidad de las técnicas de inteligencia artificial *versus* técnicas estadísticas clásicas.

Autora: Alba María Arceo Vilas

---

Tesis doctoral UDC 2020

Directores: Salvador Pita Fernández

Sonia Pértega Díaz

Alejandro Pazos Sierra

Tutor: Alejandro Pazos Sierra

Programa oficial de doctorado en Tecnologías de la Información y las Comunicaciones



UNIVERSIDADE DA CORUÑA



*A mi padre,  
siempre conmigo*



## *Agradecimientos*

En primer lugar, quiero agradecer al laboratorio de RNASA-IMEDIR la oportunidad que me han brindado de realizar con ellos esta Tesis Doctoral; en especial a mi director el Dr. Alejandro Pazos Sierra; por su apoyo y dedicación durante todo este proceso.

Este agradecimiento también va dirigido a la Unidad de Apoyo a la Investigación del Complejo Hospitalario Universitario de A Coruña, en concreto a mi directora la Dra. Sonia Pértega Díaz, por su tiempo y entrega.

Agradecerle a Carlos Fernández Lozano y a María Teresa García Rodríguez su paciencia, profesionalidad, y todo lo que me han enseñado y ayudado durante este tiempo. Carlos, gracias por adentrarme en el gran mundo de la inteligencia artificial.

Por supuesto quiero dar las gracias a los participantes que han formado parte de este proyecto, ya que sin ellos no se hubiera podido llevar a cabo.

A mi familia y amigos, por estar siempre a mi lado cuando los he necesitado.

A mi marido y a mi hija por todo el tiempo que no he disfrutado con ellos, gracias Noel por estar incondicionalmente conmigo todos estos años.

Y finalmente me gustaría hacer una mención especial a una gran persona con la que he comenzado este camino y no pudo ver el resultado final. A ti, Dr. Salvador Pita Fernández. Quiero agradecerte tu perseverancia y apoyo en los momentos más difíciles. No se puede expresar con palabras todo lo que has aportado a la investigación. Gracias de verdad por una vida dedicada al conocimiento.



## Resumo

Analizar o estado nutricional e a adhesión á dieta mediterránea pode servir para determinar estados de vulnerabilidade nutricional e identificar patoloxías asociadas como poden ser trastornos alimenticios ou psicolóxicos.

Polo que esta Tese de doutoramento céntrase por una banda, na recollida de información mediante entrevista persoal, exploración física, revisión das historias clínicas de atención primaria, así como na cumplimentación de diferentes cuestionarios: Mini Nutritional Assessment (MNA), índice de comorbilidade de Charlson, test de adhesión á dieta mediterránea, percepción subxectiva de peso, e as Subescalas EDI2 (insatisfacción corporal e obsesión pola delgadeza).

Por outra banda, analizarase todas estas variables heteroxéneas seguindo un deseño experimental clínico baseado en test estatísticos e se comparan os resultados con diferentes aproximacións de Intelixencia Artificial (en adiante IA), como o Machine Learning (en adiante ML) podendo despois discutir as diferencias entre ambas. A través desta comparación compróbase se se chegan ás mesmas conclusións a través das diferentes técnicas.

Este traballo dispón de autorización do Comité Ético de Investigación Clínica de Galicia (CEIC Galicia 2008/264 e 2012/390).



## **Resumen**

Analizar el estado nutricional y la adherencia a la dieta mediterránea puede servir para determinar estados de vulnerabilidad nutricional e identificar patologías asociadas como pueden ser trastornos alimenticios o psicológicos.

Por lo que esta Tesis Doctoral se centra, por un lado, en la recogida de información mediante entrevista personal, exploración física, revisión de las historias clínicas de atención primaria, así como en la cumplimentación de diferentes cuestionarios: Mini Nutritional Assessment (MNA), índice de comorbilidad de Charlson, test de adherencia a la dieta mediterránea, percepción subjetiva de peso, y las Subescalas del EDI2 (insatisfacción corporal y obsesión por la delgadez).

Por otro lado, se analizan todas estas variables heterogéneas siguiendo un diseño experimental clínico basado en test estadísticos y se comparan los resultados con diferentes aproximaciones de Inteligencia Artificial (en adelante IA), como el Machine Learning (en adelante, ML) pudiendo después discutir las diferencias entre ambas. A través de ésta comparativa se comprueba si se llega a las mismas conclusiones a través de las diferentes técnicas.

Este trabajo dispone de autorización del Comité Ético de Investigación Clínica de Galicia (CEIC Galicia 2008/264 y 2012/390).



## **Abstract**

Analyzing the nutritional status and adherence to the Mediterranean diet can be used to determine states of nutritional vulnerability and identify associated pathologies such as eating or psychological disorders.

So this Doctoral Thesis focuses, on the one hand, on the collection of information through personal interview, physical examination, review of the primary care clinical records, as well as on the completion of different questionnaires: Mini Nutritional Assessment (MNA), Charlson comorbidity index, test of adherence to the Mediterranean diet, subjective perception of weight, and EDI2 subscales (body dissatisfaction and obsession with thinness).

On the other hand, all these heterogeneous variables are analyzed following a clinical experimental design based on statistical tests and the results are compared with different Artificial Intelligence (hereinafter, IA) approaches, such as Machine Learning (hereinafter, ML) and can then discuss the differences between them. Through this comparison, it is checked whether the same conclusions are reached through the different techniques.

This work is authorized by the Clinical Research Ethics Committee of Galicia (CEIC Galicia 2008/264 and 2012/390).



## **Prefacio**

Una temprana valoración nutricional permite evitar, o al menos paliar, los posibles riesgos de salud asociados a una mala alimentación.

La malnutrición es considerada una mala nutrición motivada por diferentes causas como ingerir poco, comer en exceso, consumir alimentos de escaso o nulo valor nutricional o mezclarlos inadecuadamente.

La presente Tesis Doctoral se ha enfocado hacia la malnutrición en exceso ya que actualmente, en los países desarrollados, el aumento de peso es la segunda causa prevenible y evitable de enfermedades crónicas no transmisibles y múltiples exitus.

Por un lado, la realización de una evaluación completa y exhaustiva del estado nutricional permite planificar y corregir las desviaciones de la normalidad controlando la comorbilidad asociada al problema en cuestión.

Por otro lado, la dieta mediterránea, declarada Patrimonio Cultural Inmaterial de la Humanidad en el 2010, posee unas cualidades que justifican y demuestran numerosos beneficios en materia de salud, entre los que se encuentra una disminución de la incidencia de eventos cardiovasculares. Estas ventajas se traducen en una mejor calidad de vida y menor riesgo de patologías crónicas, por lo que es de gran interés ver el grado de adherencia a dicha dieta entre la comunidad.

Para analizar y conocer los resultados poblacionales de este trabajo se han empleado técnicas estadísticas clásicas, y técnicas de IA como el ML, lo que permitirá comparar la bondad de ambas técnicas para analizar este tipo de datos, que permiten obtener conclusiones de la investigación empírica mediante el uso de modelos matemáticos y de IA.

Las técnicas de IA, entre las que se encuentra el ML, proporcionan a las computadoras la capacidad de aprender y de manejar grandes cantidades de datos e informaciones para generar nuevos conocimientos. El aprendizaje por parte de la computadora se realiza de forma “socrática”; esto es, a partir de ejemplos y no a partir de instrucciones. Para ello, se le han de proporcionar un conjunto completo de ejemplos (pares de entrada–salida deseada), incluyendo contraejemplos, y el sistema de IA será

capaz, a partir de dicho conjunto de ejemplos, de representar de forma distribuida entre las conexiones de sus elementos, el conocimiento que se le aporta. Esta forma de representación distribuida del conocimiento le aportará al sistema la capacidad de seguir funcionando adecuadamente en entornos del mundo real, dónde estará sometido a fallos en sus elementos y a ruidos del entorno. De esta manera, una vez entrenado adecuadamente el sistema, será capaz de que cuando se le presente una nueva entrada (incluyendo las que se usaron en el proceso de entrenamiento y nuevas entradas con las que previamente no haya tenido contacto) el sistema será capaz de ofrecer como respuesta la salida correcta. Esto aporta una eficiente ayuda en la toma de decisiones facilitando el diagnóstico clínico, el tratamiento, el control y seguimiento de los distintos procesos clínicos que se plantean.

El objetivo principal de esta Tesis es establecer, o dar a conocer, nuevos factores de riesgo asociados a una mala nutrición y a una baja adherencia a la dieta mediterránea que conlleven un aumento de prevalencia de enfermedades relacionadas con una mala alimentación, por lo que se ha realizado un estudio comparativo entre las técnicas estadísticas clásicas y los métodos de ML, discutiendo posteriormente los resultados a los que llega cada método.

<b>1. INTRODUCCIÓN Y JUSTIFICACIÓN DEL ESTUDIO.....</b>	<b>25</b>
<b>1.1. ESTADO NUTRICIONAL: SOBREPESO Y OBESIDAD EN POBLACIÓN ADULTA .....</b>	<b>27</b>
1.1.1. COSTES DIRECTOS E INDIRECTOS DEL EXCESO DE PESO.....	28
1.1.2. VALORACIÓN DEL ESTADO NUTRICIONAL.....	29
1.1.3. PERCEPCIÓN DE LA IMAGEN CORPORAL .....	32
1.1.4. EDUCACIÓN SANITARIA COMO PAPEL FUNDAMENTAL DE LA ENFERMERÍA .....	34
<b>1.2. DIETA MEDITERRÁNEA .....</b>	<b>36</b>
1.2.1. COMPUESTOS BIOACTIVOS DE LA DIETA MEDITERRÁNEA.....	37
1.2.2. ESTADO DE LA DIETA ESPAÑOLA.....	39
1.2.3. EVIDENCIA ACUMULADA ASOCIADA A LA DIETA MEDITERRÁNEA .....	40
<b>1.3. PAPEL DE LAS TÉCNICAS DE INTELIGENCIA ARTIFICIAL EN LA INVESTIGACIÓN BIOMÉDICA .....</b>	<b>43</b>
1.3.1. TÉCNICAS DE INTELIGENCIA ARTIFICIAL Y APLICABILIDAD EN BIOMEDICINA .....	44
<b>1.4. JUSTIFICACIÓN DEL ESTUDIO.....</b>	<b>48</b>
<b>2. HIPÓTESIS Y OBJETIVOS .....</b>	<b>49</b>
<b>3. METODOLOGÍA, MATERIAL Y FASES.....</b>	<b>53</b>
<b>3.1 METODOLOGÍA.....</b>	<b>55</b>
3.1.1. DISEÑO DEL ESTUDIO .....	55
3.1.2. CRITERIOS DE INCLUSIÓN Y EXCLUSIÓN.....	56
3.1.3. JUSTIFICACIÓN DEL TAMAÑO MUESTRAL.....	56
3.1.4. MEDICIONES .....	57
3.1.4.1. Variables de identificación del paciente.....	57
3.1.4.2. Variables sociodemográficas.....	57
3.1.4.3. Comorbilidad.....	58
3.1.4.4. Variables antropométricas.....	60
3.1.4.5. Valoración del estado nutricional.....	64
3.1.4.6. Test de adherencia a la dieta mediterránea.....	64

3.1.4.7	Autopercepción de la imagen corporal.....	65
3.1.5	ASPECTOS ÉTICO-LEGALES.....	66
3.1.6	ANÁLISIS DE LOS DATOS OBTENIDOS EN EL TRABAJO DE CAMPO.....	67
3.1.6.1.	Técnicas “clásicas” de análisis estadístico .....	67
3.1.6.2	Técnicas de inteligencia artificial (Machine Learning).....	67
<b>3.2</b>	<b>MATERIAL Y FASES .....</b>	<b>69</b>
3.2.1	FASES COMUNES DEL DISEÑO EXPERIMENTAL .....	69
3.2.1.1	Selección de datos .....	69
3.2.1.2.	Preprocesado.....	70
3.2.1.2.1.	<i>Preparación de los datos</i> .....	70
3.2.2	FASES ESPECÍFICAS DEL DISEÑO EXPERIMENTAL DE ML.....	71
3.2.2.1.	Normalización .....	71
3.2.2.2	Aprendizaje.....	71
3.2.2.2.1.	<i>Entrenamiento del modelo</i> .....	72
3.2.2.2.2	<i>Reducción de la dimensionalidad</i> .....	73
3.2.2.2.3	<i>Estimación del error real del modelo</i> .....	74
3.2.2.3.	Selección del mejor modelo .....	77
3.2.2.4	Técnicas de Machine Learning.....	79
3.2.2.4.1	<i>Máquinas de soporte vectorial</i> .....	79
3.2.2.4.2	<i>Random Forest</i> .....	84
3.2.2.4.3	<i>K vecinos más próximos (kNN)</i> .....	85
3.2.2.4.4	<i>Modelos lineales generalizados</i> .....	89
<b>4.</b>	<b>RESULTADOS .....</b>	<b>95</b>
4.1.	CARACTERÍSTICAS GENERALES DE LA MUESTRA ESTUDIADA ..	97
4.1.1.	Variables sociodemográficas.....	97
4.1.2.	Comorbilidad y hábito tabáquico .....	99
4.1.3.	Características antropométricas.....	102
4.1.4.	Autopercepción de la imagen corporal.....	105
4.1.4.1	Percepción subjetiva del peso.....	105
4.1.4.2	Imagen corporal .....	106
4.2.	EVALUACIÓN DEL ESTADO NUTRICIONAL.....	108

4.2.1. Prevalencia y riesgo de malnutrición según el Mini Nutritional Assessment (MNA).....	108
4.2.2. Factores asociados al riesgo de malnutrición.....	112
4.2.2.1    Técnicas estadísticas.....	112
4.2.2.2    Técnicas de inteligencia artificial.....	117
<b>4.3. ADHERENCIA A LA DIETA MEDITERRÁNEA.....</b>	<b>119</b>
4.3.1. Grado de adherencia a la dieta mediterránea .....	119
4.3.2. Factores asociados a la adherencia a la dieta mediterránea .....	123
4.3.2.1    Técnicas estadísticas .....	123
4.3.2.2    Técnicas de inteligencia artificial .....	129
<b>5. DISCUSIÓN Y CONCLUSIONES .....</b>	<b>137</b>
<b>6. FUTUROS DESARROLLOS.....</b>	<b>141</b>
<b>REFERENCIAS .....</b>	<b>145</b>
<b>DIFUSIÓN DE RESULTADOS.....</b>	<b>159</b>
<b>ANEXOS .....</b>	<b>161</b>
<b>ANEXO 1. HOJA DE INFORMACIÓN AL PACIENTE .....</b>	<b>162</b>
<b>ANEXO 2. TRÍPTICO INFORMATIVO DEL ESTUDIO (HOJA 1) .....</b>	<b>164</b>
<b>ANEXO 2. TRÍPTICO INFORMATIVO DEL ESTUDIO (HOJA 2) .....</b>	<b>165</b>
<b>ANEXO 3. CONSENTIMIENTO INFORMADO DEL PACIENTE .....</b>	<b>166</b>
<b>ANEXO 4. CUADERNO DE RECOGIDA DE DATOS (40-64 AÑOS).....</b>	<b>167</b>
<b>ANEXO 5. CUADERNO DE RECOGIDA DE DATOS (65 Y MÁS AÑOS) ...</b>	<b>170</b>
<b>ANEXO 6. SCORE DE COMORBILIDAD DE CHARLSON.....</b>	<b>173</b>
<b>ANEXO 7. TEST DE FAGERSTRÖM.....</b>	<b>174</b>
<b>ANEXO 8. CUESTIONARIO MNA (MINI NUTRITIONAL ASSESSMENT).....</b>	<b>176</b>
<b>ANEXO 9. TEST DE ADHERENCIA A LA DIETA MEDITERRÁNEA .....</b>	<b>177</b>
<b>ANEXO 10. CUESTIONARIO EATING DISORDER INVENTORY (EDI-IC Y EDI-OD).....</b>	<b>178</b>
<b>ANEXO 11. CEIC DE GALICIA 2008.....</b>	<b>180</b>
<b>ANEXO 12. CEIC DE GALICIA 2012.....</b>	<b>181</b>
<b>ÍNDICE DE PALABRAS SIGNIFICATIVAS .....</b>	<b>182</b>

## TABLAS

Tabla 1. <i>Cuestionarios sobre la percepción de la imagen corporal</i> .....	33
Tabla 2. <i>Factores alimenticios y frecuencia de ingesta de la Dieta Mediterránea</i> .....	37
Tabla 3. <i>Efectos de la Dieta Mediterránea en la enfermedad cardiovascular</i> .....	42
Tabla 4. <i>Datos poblacionales del municipio de Cambre (A Coruña) del año 2012 y datos muestrales según edad y sexo.</i> ....	57
Tabla 5. <i>Grado de comorbilidad según el Score de Charlson</i> .....	58
Tabla 6. <i>Dependencia a la nicotina según el Test de Fageström</i> .....	59
Tabla 7. <i>Clasificación de la obesidad según la OMS</i> .....	60
Tabla 8. <i>Puntos de corte del perímetro de cintura (OMS)</i> .....	61
Tabla 9. <i>Clasificación del estado nutricional según el cuestionario Mini Nutritional Assessment (MNA)</i> .....	64
Tabla 10. <i>Tipos de error en el contraste de hipótesis</i> .....	78
Tabla 11. <i>Distribución de la muestra estudiada según variables sociodemográficas.</i> ..	97
Tabla 12. <i>Características sociodemográficas de la muestra estudiada, según grupos de edad.</i> .....	98
Tabla 13. <i>Comorbilidad, tensión arterial y hábito tabáquico.</i> .....	99
Tabla 14. <i>Comorbilidad y hábito tabáquico según grupos de edad</i> .....	101
Tabla 15. <i>Características antropométricas de la muestra estudiada</i> .....	103
Tabla 16. <i>Características antropométricas según grupos de edad</i> .....	104
Tabla 17. <i>Percepción subjetiva de peso en la muestra estudiada</i> .....	105
Tabla 18. <i>Percepción subjetiva de peso según grupos de edad</i> .....	106
Tabla 19. <i>Distribución de la muestra total estudiada según las subescalas EDI-IC y EDI-OD</i> .....	106
Tabla 20. <i>Distribución de la muestra estudiada según las subescalas EDI-IC y EDI-OD en los grupos de edad.</i> .....	108
Tabla 21. <i>Distribución de la muestra estudiada según la evaluación del estado nutricional.</i> .....	109
Tabla 22. <i>Puntuación global del MNA y prevalencia de malnutrición y riesgo de malnutrición en la muestra total.</i> .....	111

Tabla 23. <i>Puntuación global del MNA y prevalencia de malnutrición y riesgo de malnutrición en los grupos de edad.</i> .....	112
Tabla 24. <i>Correlación de la puntuación MNA con diferentes variables.</i> .....	113
Tabla 25. <i>Puntuación media del MNA según diferentes variables de interés</i> .....	114
Tabla 26. <i>Análisis bivariado para determinar variables asociadas a la presencia de malnutrición o riesgo de malnutrición según el MNA.</i> .....	115
Tabla 27. <i>Análisis multivariado de regresión logística para determinar variables asociadas a la presencia de malnutrición o riesgo de malnutrición según el MNA.</i> ...	117
Tabla 28. <i>Distribución de la muestra estudiada según la adherencia a la dieta mediterránea</i> .....	119
Tabla 29. <i>Puntuación global y categórica según la adherencia a la dieta mediterránea en la muestra total</i> .....	122
Tabla 30. <i>Puntuación global y categórica según la adherencia a la dieta mediterránea en los grupos de edad</i> .....	122
Tabla 31. <i>Correlación de la puntuación del cuestionario de adherencia a la dieta mediterránea con diferentes variables</i> .....	123
Tabla 32. <i>Puntuación media del cuestionario de adherencia a la dieta mediterránea según diferentes variables de interés</i> .....	125
Tabla 33. <i>Análisis bivariado para determinar variables cuantitativas asociadas a la buena o baja adherencia a la dieta mediterránea</i> .....	126
Tabla 34. <i>Análisis bivariado para determinar variables cualitativas asociadas a la buena o baja adherencia a la dieta mediterránea</i> .....	127
Tabla 35. <i>Análisis multivariante de regresión logística para determinar variables asociadas de forma independiente a una baja adherencia a la dieta mediterránea.</i> ...	129
Tabla 36. <i>Clasificación promedio obtenida con la prueba de Friedman y la extensión de Iman y Davenport.</i> .....	134

## FIGURAS

<i>Figura 1.</i> Conducta alimentaria.....	35
<i>Figura 2.</i> Localización del municipio de Cambre.....	55
<i>Figura 3.</i> Validación cruzada de “k iteraciones” .....	72
<i>Figura 4.</i> Matriz de confusión.....	74
<i>Figura 5.</i> Curva ROC perfecta .....	75
<i>Figura 6.</i> AUROC sin valor diagnóstico.....	76
<i>Figura 7.</i> Hiperplano óptimo de separación.....	80
<i>Figura 8.</i> Clasificación del patrón X con 1NN .....	86
<i>Figura 9.</i> Distribución de pacientes según la puntuación total de la subescala insatisfacción corporal (EDI-IC) .....	107
<i>Figura 10.</i> Distribución de pacientes según la puntuación total de la subescala obsesión por la delgadez (EDI-OD) .....	107
<i>Figura 11.</i> Histograma de puntuación total del MNA (Mini Nutritional Assessment) .....	111
<i>Figura 12.</i> “Boxplots” de resultados de los algoritmos de aprendizaje automático.....	118
<i>Figura 13.</i> Distribución de las puntuaciones del cuestionario de adherencia a la dieta mediterránea. ....	121
<i>Figura 14.</i> Distribución de pacientes según la adherencia a la dieta mediterránea.....	121
<i>Figura 15.</i> Resumen del desempeño. AUROC clasificación agregada de validación cruzada.....	130
<i>Figura 16.</i> Estabilidad de los algoritmos de aprendizaje automático durante los experimentos.....	131
<i>Figura 17.</i> Gráfica Q-Q de valores observados <i>versus</i> valores esperados. ....	132
<i>Figura 18.</i> Gráfica Q-Q de los valores observados <i>versus</i> los esperados agrupados por conjunto de datos .....	133
<i>Figura 19.</i> Importancia de la característica según el parámetro $w$ del modelo ganador. ....	135

## LISTA DE ABREVIATURAS

<b>AGB</b>	Área Grasa del Brazo
<b>AMB</b>	Área Muscular del Brazo
<b>AUC</b>	Área bajo la curva
<b>BIA</b>	Bioimpedancia Eléctrica
<b>CMBD</b>	Conjunto Mínimo Básico de Datos
<b>CONUT</b>	Sistema de Control Nutricional (CONtrolling NUTricional status)
<b>DMed</b>	Dieta mediterránea
<b>ECV</b>	Enfermedad CardioVascular
<b>ENPE</b>	Estudio Nutricional de la Población Española
<b>EPIC</b>	European Prospective Investigation into Cancer and Nutrition Study
<b>ESPEN</b>	European Society of Parenteral and Enteral Nutrition
<b>EUFIC</b>	Consejo Europeo de Información sobre Alimentación
<b>IA</b>	Inteligencia artificial
<b>IAM</b>	Índice Adiposo Muscular
<b>ICC</b>	Índice de cintura-cadera
<b>IMC</b>	Índice de Masa Corporal
<b>kNN</b>	k-Nearest Neighbor
<b>ML</b>	Machine Learning
<b>MNA</b>	Mini Nutritional Assessment
<b>NRS</b>	Nutritional Risk Screening
<b>OMS</b>	Organización Mundial de la Salud
<b>PMB</b>	Perímetro Muscular del Brazo
<b>PREDIMED</b>	PREvención con Dieta MEDiterránea
<b>SEEDO</b>	Sociedad Española para el Estudio de la Obesidad
<b>SENPE</b>	Sociedad Española de Nutrición Parenteral y Enteral
<b>SUN</b>	Seguimiento Universidad de Navarra
<b>SVM</b>	Máquinas de Soporte Vectorial
<b>VSG</b>	Valoración Subjetiva Global



# 1. INTRODUCCIÓN Y JUSTIFICACIÓN DEL ESTUDIO



## 1.1. ESTADO NUTRICIONAL: SOBREPESO Y OBESIDAD EN POBLACIÓN ADULTA

El estado nutricional es uno de los indicadores esenciales de salud poblacional, siendo de gran importancia llevar a cabo una evaluación integral que identifique los grupos de riesgo y, con una apropiada interpretación de los resultados, corregir las desviaciones de la normalidad para mejorar el estado de salud, disminuyendo con ello la morbimortalidad en la población.

El término “malnutrición” posee dos vertientes: una alteración por déficit que acarrea una desnutrición y una ingesta calórica excesiva que produce obesidad; siendo ambas un desequilibrio entre los requisitos corporales y el consumo nutricional que realiza cada individuo (1,2).

Actualmente, después del tabaco, el sobrepeso expresado como Índice de Masa Corporal (en adelante, IMC)  $\geq 25 \text{ kg/m}^2$  y la obesidad (IMC  $\geq 30 \text{ kg/m}^2$ ) son la segunda causa prevenible y eludible de enfermedad y de muerte en los países desarrollados (3,4).

Según la Organización Mundial de la Salud (OMS), entre 1975 y 2016, la prevalencia mundial de obesidad casi ha triplicado su valor (5). Más de la mitad de la población mundial reside en países donde el sobrepeso y la obesidad tienen grandes consecuencias negativas para la salud de la población, y cuyo costo principal, en materia de salud, es el aumento de patologías crónicas no transmisibles derivadas de la mala alimentación, y muy prevalentes, como la hipertensión o la diabetes, entre otras (6).

Se estima que, si no se corrige a corto plazo y continúa creciendo al ritmo actual, para el año 2030 más del 40% de la población mundial padecerá sobrepeso y más de la quinta parte será obesa (7). Estados Unidos es, con diferencia, el país con mayor porcentaje de obesos entre su población adulta (38,2%).

En la Unión Europea, más de 13 millones de fallecimientos al año están relacionadas con el exceso de peso. Los países europeos con menor prevalencia son Suiza, Dinamarca e Italia, donde el porcentaje de personas con obesidad no supera el 20%. En situación opuesta se encuentran Turquía o Reino Unido, que destacan por ser los que registran un mayor porcentaje de personas con obesidad. En Inglaterra el 27,80% de su población tiene sobrepeso. Pero es Turquía el país que concentra un mayor porcentaje de personas obesas, un total del 32,10% de su población (4).

En nuestro país, según los datos de la Encuesta Nacional de Salud en España 2017, un 18,2% de los varones mayores de 17 años y un 16,7% de las mujeres padecen obesidad; y un 44,3% de hombres y un 30,0% de mujeres sufren sobrepeso (8). Este hecho

probablemente explique que la obesidad sea el primer factor de riesgo cardiovascular en España.

También en territorio nacional se debe destacar el Estudio Nutricional de la Población Española (ENPE), un proyecto transversal de población no institucionalizada realizado en 2014-15, con 3966 individuos de 25 a 64 años, donde se concluye que la prevalencia general de sobrepeso fue del 39,3% y de obesidad del 21,6% (9).

Visto por Comunidades Autónomas, en este mismo estudio se muestra que Galicia, con un 24,9% de su población adulta, y Andalucía, con un 24,4%, siguen de cerca a Asturias, completando el podio de las tres comunidades autónomas de España con la mayor prevalencia de obesos (9).

Por debajo, en orden descendente, aparecen Murcia (23,9%), Madrid (23,8%), Aragón (23,7%), Castilla - La Mancha (23,6%), Castilla y León (22,1%), Canarias (20,1%), Navarra (20,9%), Comunidad Valenciana (19,8%) y Extremadura (19%). Las comunidades autónomas que presentan una menor prevalencia de obesos entre población adulta son: La Rioja (17,9%), Cantabria (17,6%), País Vasco (16,8%), Cataluña (15,5%) y, finalmente Las Islas Baleares con un 10,5% de obesos adultos (9).

---

### 1.1.1. COSTES DIRECTOS E INDIRECTOS DEL EXCESO DE PESO

La mayoría de investigaciones poblacionales confirman que, a medida que aumenta el peso hasta alcanzar el grado de sobrepeso y obesidad, aumentan paralela y significativamente los riesgos de las siguientes afecciones (10,11):

- Diabetes mellitus: En el caso de las personas obesas, el riesgo de padecer diabetes es tres veces mayor al de las personas que no tienen obesidad. Según la OMS, la prevalencia mundial de diabetes ha pasado de un 4,7% en 1980 a casi el doble, un 8,5%, en 2014 (12), destacando que el 80 % de las personas diabéticas fallecen a causa de complicaciones cardiovasculares.
- Hipertensión arterial (presión arterial sistólica y/o diastólica igual o superior a 140/90 mm Hg): La causa principal de hipertensión arterial es la obesidad, además de ser el factor primordial que entorpece y complica su control (13).
- Dislipemia: La dislipidemia en la obesidad se caracteriza principalmente por una hipertrigliceridemia, causada por el exceso de ácidos grasos libres procedentes del tejido adiposo. La lipólisis de las lipoproteínas en la obesidad está disminuida lo que provoca un aumento de los niveles de triglicéridos en sangre. Como consecuencia, las lipoproteínas HDL y LDL transportadoras de colesterol,

muestran morfología y composición anómala asociadas con un mayor riesgo cardiovascular.

- Patología cardíaca: provocada por cardiopatía coronaria, insuficiencia cardíaca y accidente cerebrovascular.
- Complicaciones óseas y articulares: un exceso de peso ejerce mayor presión sobre los huesos y articulaciones provocando osteoartritis (enfermedad que origina rigidez y dolor articular) o artrosis (degeneración del cartílago y el hueso subyacente de una articulación).
- Apnea del sueño, pudiendo derivar en fatiga, mala atención y somnolencia diurna.
- Algunos tipos de tumores, como cáncer de endometrio, mama o colon.
- Patología hepática y cálculos biliares.
- Patología ginecológica, como una menstruación anómala o infertilidad.

A nivel internacional, los estudios que analizan las cuantías económicas derivadas de la obesidad y el sobrepeso han confirmado a través del Consejo Europeo de Información sobre Alimentación (EUFIC), que representan entre un 2% y un 7% del total del gasto sanitario (un 7% en España; unos 2.500 millones de euros al año), pudiendo llegar a más de un 12% en Estados Unidos (14).

A mayores, si se tienen en cuenta los costes que conllevan las enfermedades relacionadas con la obesidad, es decir, la comorbilidad, son 900 millones más, procedentes de enfermedades cardiovasculares (22,6%), diabetes (9%), dislipemia (2,4%), y enfermedades musculoesqueléticas (1%) (14).

A los importes comentados anteriormente hay que añadir el coste social que provoca la disminución de la productividad laboral, ya que las personas obesas registran menos tasas de participación, y por lo tanto menores sueldos. Las bajas laborales o incapacidades suponen 1.600 millones de euros, 18 millones de días de baja y 30.000 defunciones cada año, con una pérdida de 40.000 años de vida laboral y una esperanza de vida nueve años menor (14).

---

### 1.1.2. VALORACIÓN DEL ESTADO NUTRICIONAL

La valoración adecuada del estado nutricional resulta esencial en la planificación de intervenciones para corregir el exceso de peso y, de esta manera, poder controlar la comorbilidad asociada.

Por ello, en función de los objetivos planteados, la evaluación nutricional es realizada a través de diferentes indicadores, formando un método diagnóstico global. Hay

gran variedad de procedimientos para determinar el estado nutricional de los pacientes; entre ellos, se encuentran:

#### **a) Antropometría:**

La antropometría es un procedimiento de fácil aplicación, económico y no invasivo por lo que ha sido ampliamente utilizada para estimar el estado nutricional a nivel clínico y, o, epidemiológico.

A través de la medición de diferentes parámetros antropométricos, se observa el nivel de reservas proteicas y calóricas, sus desequilibrios y posibles consecuencias.

Las mediciones antropométricas más importantes y utilizadas, debido a su sencillez en la determinación, son:

- El IMC: es la relación entre el peso del sujeto (en kilogramos) y el cuadrado de su talla (indicada en metros).
- El índice de cintura-cadera (ICC): es la relación entre el perímetro de cintura y el perímetro de cadera (ambos en centímetros).
- Los pliegues cutáneos: medidos con un plicómetro en distintas zonas corporales. Siendo las más frecuentes la zona tricípital, bicípital, subescapular y suprailíaca (expresados en milímetros).
- El perímetro de brazo: junto con el pliegue cutáneo tricípital, través de fórmulas establecidas previamente, permite obtener el Perímetro Muscular del Brazo (PMB), el Área Muscular del Brazo (AMB), el Área Grasa del Brazo (AGB) y el Índice Adiposo Muscular (IAM).

#### **b) Estudios de composición corporal mediante métodos bioeléctricos:**

A finales de los 80 empezaron a surgir los primeros aparatos de Bioimpedancia Eléctrica (BIA). Debido a la simplicidad de su protocolo no invasivo, rapidez de administración y relativa economía, estos aparatos han tenido una gran aceptación.

Los estudios de BIA se basan en la relación que hay entre las propiedades eléctricas del cuerpo humano, la composición de los tejidos y del contenido total de agua corporal, permitiendo diferenciar qué proporción de peso se corresponde con tejido magro (musculatura), y que parte está formada por tejido graso, entre otros.

Se considera una masa adiposa excesiva  $> 25\%$  en varones y  $> 33\%$  en mujeres (15).

### c) Cuestionarios nutricionales:

#### **Valoración Subjetiva Global (VSG)**

Esta valoración clasifica a los individuos de manera subjetiva en función de datos recogidos a través de su historial clínico (cinco preguntas sobre: la pérdida ponderal, ingesta actual, síntomas gastrointestinales, capacidad funcional y demandas metabólicas) y la exploración física (cuatro valoraciones: pérdida de grasa subcutánea en tríceps y tórax; pérdida de músculo en cuádriceps y deltoides; edema en tobillo y sacro; y presencia de ascitis) (16).

#### **Nutritional Risk Screening (NRS-2002)**

Este test es empleado para averiguar el riesgo nutricional en pacientes hospitalizados. Fue planteado en el 2003 por la *European Society of Parenteral and Enteral Nutrition* (ESPEN) (17). Se compone de dos apartados, un cribado inicial de cuatro preguntas, y si hay una respuesta afirmativa, incluso en una sola de ellas, se sigue con la segunda parte, una valoración nutricional sobre el descenso de peso, reducción de la ingesta e IMC.

También se recopila información sobre la gravedad de las posibles patologías de los participantes. En los  $\geq 70$  años se añadirá 1 punto más por edad.

Si la puntuación obtenida es mayor o igual a 3 indica que el sujeto se encuentra en riesgo nutricional, por lo que se debe comenzar un tratamiento nutricional. En caso de que sea menor de 3, se reevaluará al paciente semanalmente.

#### **Mini Nutritional Assessment (MNA)**

La principal aportación del MNA es facilitar de manera fiable, sencilla y sin riesgos, la realización de la valoración nutricional. Además, permite la detección del riesgo nutricional sin hacer otras pruebas complementarias, teniendo valor propio como pronóstico de supervivencia.

Engloba indicadores antropométricos, dietéticos, de evaluación global y de autopercepción de salud y nutrición. Esta herramienta tiene una elevada sensibilidad (96%) y especificidad (98%), siendo válido su resultado en distintos ámbitos asistenciales (18).

Consta de 18 ítems y clasifica al paciente en tres grupos: estado nutricional normal, riesgo de malnutrición y malnutrición. Permite realizar un diagnóstico nutricional, a nivel individual o colectivo, catalogando la malnutrición desde el punto de vista cuantitativo en sus diferentes grados de severidad, siguiendo las instrucciones de la guía de cumplimentación (18).

#### **d) Parámetros analíticos:**

Entre los parámetros analíticos más utilizados están la albúmina, la prealbúmina, las proteínas totales y la transferrina. Uno de los métodos de valoración nutricional que utiliza estos parámetros es el Sistema de Control Nutricional (CONUT). Este sistema (19) es un método informático desarrollado por la Unidad de Nutrición del Hospital de la Princesa de Madrid y respaldado por la Sociedad Española de Nutrición Parenteral y Enteral (SENPE) con el que se detecta precozmente la malnutrición. Está basado en el empleo sistemático de datos demográficos del Conjunto Mínimo Básico de Datos (CMBD), otros datos relacionados con su estado de salud (Farmacia, Nutrición y Dietética, ...) y analíticos (albúmina sérica, colesterol y linfocitos) de los participantes.

En cuanto a los estudios nutricionales más representativos a nivel nacional, cabe destacar que la gran mayoría están destinados a mayores de 65 años, debido a la prevalencia de malnutrición y la importancia de la detección precoz del riesgo nutricional en esta etapa vital, especialmente vulnerable.

La mayor prevalencia registrada a nivel nacional de riesgo nutricional en adultos mayores sin institucionalizar (57.5%) se detectó en Orense, en 2011. Mediante un estudio de 728 personas autónomas de 65 años que viven en sus domicilios (20). En otros estudios nacionales, el riesgo de malnutrición se encuentra entre un 23.9% y un 31.5% (21–23).

---

#### 1.1.3. PERCEPCIÓN DE LA IMAGEN CORPORAL

El término de imagen corporal abarca tanto a la idea que se posee de la totalidad del cuerpo y de cada una de sus zonas, como del movimiento y límites de éste. En las investigaciones sobre las alteraciones de la conducta alimentaria, se considera una determinación cualitativa de la situación nutritiva del sujeto (24).

Existen factores destacables como: sobrepeso u obesidad, género masculino, escasos años de formación reglada, pertenecientes a la tercera edad y grupo socioeconómico bajo, los cuales tienen una menor propensión a considerarse con sobrepeso (25).

La percepción que la persona tiene de su propia imagen corporal es determinante para su peso. La estructura de la imagen corporal es cambiante, ya que la variación de algún elemento provoca la modificación del resto. Con el paso de los años, la imagen puede variar hacia una considerable aceptación, o, por lo contrario, una repulsa total.

En la década de los 80, los estudios analizaban los factores perceptivos de la imagen corporal; sin embargo, en los años 90, la valoración se ha ido desviando hacia los intrínsecos y actitudinales.

Para la evaluación de la imagen corporal, Rosen et al. (26) considera las siguientes áreas:

- Apreciación de la forma y tamaño del cuerpo.
- Áreas cognitivas y emocionales: creencias y preocupaciones, la molestia que causan y el nivel de convencimiento sobre la presencia del defecto, así como el empeño que realiza el individuo para dominar la idea.
- Áreas conductuales: evasión, ocultamiento, ritos, etc.
- Otras alteraciones psicológicas: depresión, alteración de la alimentación, estrés social, etc.
- Rasgos de personalidad: autoestima, habilidades sociales, niveles de adaptación, etc.

Los cuestionarios más empleados para determinar la percepción de la imagen corporal desde las distintas perspectivas se muestran en la Tabla 1 (26).

Tabla 1. *Cuestionarios sobre la percepción de la imagen corporal*

CUESTIONARIOS PERCEPCIÓN CORPORAL			
TIPO	CUESTIONARIO	AUTOR/ES	AÑO
GENERALES	Subescalas del Eating Disorders Inventory-2 (EDI-2)(27)	Garner, Olmstead y Polivy	1983
	Escala Yale-Brown modificada (28)	Phillips	1993
	Body shape questionnaire (BSQ) (29)	Cooper, Taylor, Cooper y Fairburn	1987
	Escala de Síntomas del Trastorno Dismórfico (30)	Peruggi, Giannotti, Di Vaio, Frare, Saettoni y Cassano	1996
COGNITIVOS	Multidimensional Body Self-Relations Questionnaire (MBSRQ) (31)	Cash, Winstead y Janda	1986
	Body Image Automatic Thoughts Questionnaire (BIATQ) (32)	Rucker y Cash	1992
	Appearance Schemas Inventory (ASI) (33)	Cash y Labarge	1996
CONDUCTUALES	Cuestionario de Evitación de la Imagen Corporal (BIAQ) (34)	Rosen, Srebnik, Saltzberg y Wendt	1991

En aquellos individuos que, tras analizar sus dimensiones corporales, expresan juicios valorativos que no concuerdan con la realidad medible se objetiva que existe un trastorno de la imagen corporal, la cual presenta una valoración corporal negativa, tanto cognitiva como actitudinal. Dicha alteración constituye un indicador decisivo para efectuar algún tipo de dieta personalizada.

Según Rosen (26), estas alteraciones son lo que principalmente diferencia a los trastornos alimenticios de otras modificaciones de la ingesta y variaciones de peso. Los sujetos con trastornos alimenticios y aquellas personas que están muy pendientes de su imagen, sufren dificultades para simbolizar, evaluar y sentir su propio cuerpo.

---

#### 1.1.4. EDUCACIÓN SANITARIA COMO PAPEL FUNDAMENTAL DE LA ENFERMERÍA

Una de las principales tareas de la Enfermería es la educación para la salud, convirtiéndose en un interlocutor esencial para asesorar, tanto en la salud como en la enfermedad, a los individuos y a la comunidad. Por ello, uno de los objetivos principales de la Enfermería es el cuidado integral al paciente, su familia y la comunidad; consiguiendo que las personas obtengan el máximo nivel de autocuidados mejorando su calidad de vida.

Es verdaderamente significativa la función de la educación nutricional, ya que la adaptación de los patrones alimentarios actuales hacia otros más sanos, tendría una gran repercusión disminuyendo la prevalencia de las enfermedades derivadas de la mala alimentación (35).

En adultos, es fundamental comenzar por la motivación individual, desmontando todas aquellas creencias, tradiciones y gustos personales para evitar el abandono del programa o la negativa a seguir las pautas recomendadas.

Los cuidados de las enfermeras se realizan desde el área “biopsicosocial” del paciente y con un desarrollo estructurado en diferentes etapas.

El personal de enfermería realiza funciones globales y específicas como:

1. Empodera a la población mejorando su nivel de conocimientos nutricionales.
2. Lleva a cabo la cadena de modificación del comportamiento alimentario (Fig. 1)



*Figura 1. Conducta alimentaria.*

3. Valora las necesidades nutricionales en los distintos ciclos vitales y períodos de enfermedad.
4. Enseña cómo debe ser una nutrición equilibrada.
5. Ayuda a conseguir un IMC adecuado a las circunstancias.
6. Realiza un seguimiento nutricional global y de manera continuada.
7. Realiza una evaluación final con los resultados obtenidos.

Acerca de la evaluación sobre la adquisición de conocimientos y su aplicación, se considera que, por las características particulares de un entorno con una elevada carga asistencial, ha de ser pautada de manera habitual y regular, pero sobre todo con una actitud científica, negociadora, empática y enérgica.

## 1.2. DIETA MEDITERRÁNEA

El consumo de una alimentación saludable como la dieta mediterránea (en adelante, DMed) ha demostrado a lo largo de muchos años y en numerosos estudios llevados a cabo internacionalmente, una clara evidencia de múltiples beneficios (36).

El término de DMed se establece como resultado del “Estudio de los Siete Países”, promovido por el Dr. Ancel Keys entre 1958 y 1964, donde se estudió la relación de la dieta y el estilo de vida con los factores de riesgo y la incidencia de enfermedades cardiovasculares, incluyendo, en especial, la muerte por infarto o ictus (37).

Los participantes fueron 12.763 sujetos, entre 40 y 59 años, seleccionados en 16 cohortes de 7 países: Estados Unidos, Japón, Finlandia, Holanda, la antigua Yugoslavia, Italia y Grecia (37).

Las conclusiones mostraron un descenso en la mortalidad por patología coronaria y una mayor esperanza de vida en los países mediterráneos, especialmente en Grecia (38); por lo que se relacionó este hecho al tipo de dieta y al estilo de vida propio de esta zona, llevando a Keys a definirla como “Dieta Mediterránea”. En 2010, fue proclamada Patrimonio Cultural Inmaterial de la Humanidad (37).

Esta dieta presenta una frecuencia de ingesta (Tabla 2) y unos aspectos comunes que abarcan (39):

- Escaso consumo de grasas saturadas y proteínas de origen animal (mantequilla, leche entera y carnes rojas).
- Elevada ingesta de grasas monoinsaturadas (sobre todo aceite de oliva).
- Antioxidantes (frutas, verduras, vino con moderación, aceite de oliva virgen, especias y hierbas).
- Notable consumo de fibra (verduras, frutas, cereales integrales, legumbres y frutos secos).
- Balance apropiado de ácidos grasos poliinsaturados omega-6 y omega-3 (pescados, mariscos y frutos secos).

Tabla 2. Factores alimenticios y frecuencia de ingesta de la Dieta Mediterránea (39)

GRUPO DE ALIMENTOS	FRECUENCIA DE CONSUMO
Verduras u hortalizas	2 o más raciones/día
Frutas	3 o más piezas/día
Aceite de oliva	2 o más cucharadas /día
Cereales	3 a 4 veces/día
Legumbres	Al menos 3 veces/semana
Frutos secos	1 o más veces/semana
Productos lácteos, yogurt y quesos	2 a 4 porciones/día
Huevos	1 a 4 unidades/ semana
Pescados y mariscos	3 o más veces/ semana
Aves	2 a 4 veces/semana
Carnes rojas y procesadas	1 vez/semana o menos
Vino	3 o más vasos/semana

### 1.2.1. COMPUESTOS BIOACTIVOS DE LA DIETA MEDITERRÁNEA

Los principales responsables de las ventajas que ofrece esta dieta son los compuestos denominados bioactivos. Entre éstos se encuentran los antioxidantes, la fibra y los fitoesteroles, los ácidos grasos monoinsaturados, los ácidos grasos omega 3 y los probióticos.

A través de ellos, la DMed produce alteraciones fisiológicas y metabólicas que se traducen en un riesgo reducido de patologías crónicas y considerablemente una mayor y mejor esperanza y calidad de vida.

Por una parte, la DMed aconseja una baja ingesta de nutrientes significativamente nocivos, como lo son las grasas saturadas e hidratos de carbono simples, relacionados con un elevado riesgo de aterosclerosis y diabetes mellitus.

Por otra parte, el menor uso de sal conlleva un consumo reducido de sodio que, unido a un mayor consumo de magnesio y potasio, precedente de frutas y verduras,

favorece la función cardiovascular disminuyendo la tensión arterial y protegiendo el endotelio.

Asimismo, el consumo de “sofrito” (aceite de oliva, cebolla o puerro, ajo y tomate) en la comida, se asocia con un menor riesgo cardiovascular (40).

#### *Antioxidantes*

La DMed contribuye con un gran aporte de antioxidantes de origen vegetal, entre ellos, vitamina E, vitamina C,  $\beta$ -caroteno, licopeno, glutatión y polifenoles, a la mejoría a nivel endotelial, lípidos séricos y a un descenso de la presión arterial. En el estudio PREDIMED (PREvención con Dieta MEDiterránea), donde se evalúan los efectos de la DMed en la prevención primaria de la enfermedad cardiovascular (en adelante, ECV), se ha visto que una mayor ingesta de polifenoles está relacionado con una reducción del 46% del riesgo de complicaciones cardiovasculares mayores (41).

#### *Fibra*

La fibra dietética es la parte nutritiva de las plantas que no es afectada por el proceso digestivo, tolerando la digestión y absorción en el intestino delgado, y posterior fermentación en el grueso. Existen dos tipos: la fibra insoluble, como la celulosa, proviene esencialmente de cereales integrales; y la fibra soluble que está presente en las verduras, cebada, fruta, legumbres y avena.

Existen estudios epidemiológicos que demuestran una reducción del riesgo cardiovascular del 9% por cada 7 g/día de fibra (42,43). Esto es debido a que la fibra reduce la absorción de grasas y azúcares, regulando el colesterol y glucosa sanguíneos.

Igualmente, la fibra alimenticia ayuda al tránsito intestinal, previniendo el estreñimiento y el cáncer digestivo. Además, es un compuesto que regula y disminuye el apetito con lo que reduce el consumo energético y facilita el control de peso (44).

#### *Fitoesteroles*

Los esteroides son sustancias vegetales que abundan en: cereales, frutos secos, legumbres y aceites vegetales. En el estudio PREDIMED, se ha demostrado una relación inversa significativa entre el consumo de fitoesteroides y niveles de colesterol LDL plasmáticos (45).

#### *Ácidos grasos monoinsaturados*

La ventaja del aceite de oliva, y su acción beneficiosa sobre la ECV, varias clases de cáncer y el deterioro cognitivo, se debe a su estructura grasa que abunda en ácidos grasos monoinsaturados, principalmente ácido oleico (70%), poliinsaturados (10-15%), en la escasa proporción de ácidos grasos saturados (5%) y en la presencia elevada de compuestos polifenólicos antioxidantes (46).

El ácido oleico disminuye el colesterol total, el LDL y los triglicéridos y eleva el HDL. Los compuestos polifenólicos antioxidantes poseen una propiedad antitrombótica, anticancerígena, antiinflamatoria vasodilatadora e hipotensora (46). Asimismo, el aceite de oliva virgen reduce la resistencia periférica a la acción de la insulina y reduce la inflamación.

#### *Ácidos grasos omega 3*

La DMed, se caracteriza por una elevada ingesta de pescados y mariscos; con una correcta relación de ácidos grasos omega 6/ omega 3, lo que le aporta un carácter antiinflamatorio mayor que otras dietas. Un elevado consumo de ácidos grasos omega 6 aumenta el riesgo de patologías crónicas como diabetes mellitus y aterosclerosis (47).

#### *Probióticos*

Los probióticos son cultivos de microorganismos vivos que, en dosis y cepas apropiadas, pueden conceder ventajas saludables. Los alimentos fermentados procedentes de la leche, como yogurt y quesos, son incluidos en la DMed, y su consumo reduce los parámetros inflamatorios relacionados con la aterosclerosis (48).

---

### 1.2.2. ESTADO DE LA DIETA ESPAÑOLA

Como ya se ha mencionado con anterioridad, durante los años 1961-1965, Grecia era el referente dietético mediterráneo; sin embargo, algunos autores mantienen que actualmente, debido a la implantación de nuevos hábitos alimenticios, se ubica exclusivamente en unas escasas poblaciones costeras.

España se acerca poco a poco a la predisposición de otros países sin dieta mediterránea, en relación al consumo de alimentos menos saludables, ingiriendo más cantidad de carne, pescado y huevos de lo recomendado, y escaseando en leche y derivados, frutas, verduras y hortalizas, cereales, patatas y legumbres (49).

Para determinar la variación de costumbres alimentarias, se emplean índices de adecuación a la Dieta Mediterránea. Knoops (50) propone *Mediterranean Diet Score*, un índice donde el resultado de la dieta española es de 4 sobre un límite de 9. Según el *Healthy Diet Indicator* planteado por la OMS, la dieta española obtiene un 2 sobre 9 (49).

En la valoración del Panel de Consumo de Alimentos se objetiva que la energía media es de 2.761 kcal/persona y día, siendo superior al 90% de la recomendación para adultos (51).

Además, la aportación a la energía total de la dieta en España es del 14% para la proteína, 40% para los lípidos, 42% para los hidratos de carbono y 4% para el alcohol;

siendo actualmente, la contribución de los lípidos mayor de lo aconsejado. En los últimos 40 años, se percibe un aumento progresivo de grasas en la dieta, también observado en los países europeos. En cuanto a los hidratos de carbono, desde 1964, su aporte ha disminuido en más de un 10% (49).

Un 11% de la energía total procede de los ácidos grasos saturados (superior al 7% recomendado). Igualmente, es superior la cuantía de los ácidos grasos poliinsaturados (8% de la energía total), pero continúan prevaleciendo los ácidos grasos monoinsaturados (18% de la energía total). Pese a ello, los índices de calidad de la grasa son adecuados (49).

En cuanto a los micronutrientes, la dieta española es rica en vitamina E y C (el aporte de vitamina C excede el 200% de las ingestas indicadas), fundamentalmente gracias al aporte de frutas, verduras y hortalizas, estando igualmente rebasado el nivel de vitamina B12 y yodo (51).

Asimismo, el aporte medio de sodio (sal) obtenido a través de la dieta es mayor al recomendado (2.955 mg/persona y día) (51).

---

### 1.2.3. EVIDENCIA ACUMULADA ASOCIADA A LA DIETA MEDITERRÁNEA

#### *Dieta mediterránea, enfermedad cardiovascular y diabetes*

En cuanto a los efectos que posee la DMed sobre el corazón y la incidencia de diabetes se ha observado que reduce el riesgo en más del 25% de los casos (52).

Uno de los principales trabajos que muestra el vínculo entre la DMed y la ECV es la cohorte griega del estudio EPIC (*European Prospective Investigation into Cancer and Nutrition Study*), el cual demuestra que, un resultado de más de 2 puntos en el grado de adherencia a dicha dieta, se corresponde con una disminución del 33% en cuanto a la mortalidad por ECV (52).

En el estudio nacional de Martínez González MA et al., tras un seguimiento de 5 años de 13.600 adultos sin enfermedad coronaria, se percibió que 2 puntos de aumento en la adherencia a la dieta también se relacionaban con una reducción del riesgo coronario, en este caso del 26% (53).

Con respecto a la diabetes, el seguimiento de la cohorte SUN (Seguimiento Universidad de Navarra) compuesto por 13.380 adultos determinó una asociación inversa entre la ingesta de la DMed y la incidencia de diabetes, reflejada en un descenso del riesgo del 35% por cada 2 puntos de incremento en la adherencia a la DMed (54).

En el estudio de Bonaccio M et al, se asoció un conjunto de pacientes diabéticos con un mayor índice de DMed, con una reducción del 37% de la mortalidad, destacando los alimentos más preventivos como la ingesta moderada de alcohol, abundancia de cereales y verduras y escasez de lácteos y carnes (55).

#### *Dieta mediterránea y cáncer*

Diversas cohortes y estudios observacionales aluden al carácter protector de la DMed sobre la incidencia y la mortalidad por cáncer; así, el seguimiento de 22.000 sujetos durante 4 años de la cohorte EPIC resultó que la adherencia a la DMed se asociaba a una disminución de 24% en la mortalidad por cáncer (56).

En la misma línea, un metaanálisis basado en datos de 8 cohortes (514.800 individuos) llegó a la conclusión que una mayor adherencia a la DMed se relacionaba con una menor incidencia y mortalidad por cáncer (57).

Como ya se ha mencionado anteriormente, los factores responsables de este rol protector son la alta ingesta de frutas, verduras, cereales integrales y aceite de oliva y, por el contrario, el bajo consumo de carnes rojas, hidratos de carbono simples y grasas saturadas. Todo ello, con un adecuado aporte de fibra, antioxidantes, relación ácidos grasos omega6/omega3 y ácidos grasos monoinsaturados. Este patrón dietético ejerce un efecto favorable contra distintos tipos de cánceres (58).

Es de destacar que el aceite de oliva ha demostrado ser un factor independiente de protección contra varios tipos de cáncer, en especial del aparato respiratorio (58).

#### *Dieta mediterránea y enfermedades neurodegenerativas*

Las patologías neurodegenerativas producen una disfunción neuronal que, progresivamente, va atrofiando estructuras del sistema nervioso central, causando fragilidad y dependencia en la población mayor adulta. Su prevalencia alcanza el 30% en mayores de 90 años ya que se incrementa exponencialmente con la edad. Dado el envejecimiento poblacional, anualmente se diagnostican 4,6 millones de nuevos casos de demencia (59).

La DMed ha verificado en múltiples estudios reducir la incidencia de patologías como Alzheimer y Parkinson, siendo beneficiosa frente al deterioro cognitivo leve relacionado con la vejez (59).

En el metaanálisis basado en 5 cohortes de Sofi F et al con un total de 133.626 individuos, se llegó a la conclusión que un incremento de dos puntos en la adherencia a

DMed se relacionaba con un 13% de disminución del riesgo de padecer alguna alteración neurodegenerativa (57).

En la Tabla 3, se resumen los efectos beneficiosos de la dieta mediterránea.

Tabla 3. *Efectos de la dieta mediterránea en la enfermedad cardiovascular*

		
<ul style="list-style-type: none"> <li>- Disminuye la incidencia de sucesos cardiovasculares.</li> <li>- Disminuye la incidencia y mortalidad por cáncer.</li> <li>- Disminuye la incidencia de enfermedades neurodegenerativas (Parkinson y Alzheimer)</li> </ul>	<ul style="list-style-type: none"> <li>- Mejora del nivel lipídico</li> <li>- Disminuye la presión arterial</li> <li>- Disminuye el efecto nocivo del tabaquismo</li> <li>- Disminuye la incidencia de obesidad</li> <li>- Relación inversa con el perímetro de cintura y la disminución de peso</li> <li>- Prevención frente a la diabetes: disminuye la incidencia y aumenta la secreción de insulina</li> </ul>	<ul style="list-style-type: none"> <li>- Efecto antiinflamatorio</li> <li>- Mejora la función endotelial</li> <li>- Disminuye la agregación plaquetaria</li> <li>- Disminuye la progresión del proceso trombótico.</li> </ul>

### 1.3. PAPEL DE LAS TÉCNICAS DE INTELIGENCIA ARTIFICIAL EN LA INVESTIGACIÓN BIOMÉDICA

Uno de los mayores retos de la medicina actual es el hecho de poder predecir qué enfermedades podría sufrir una persona según sus hábitos (nutricionales, tóxicos, saludables...), su entorno y las condiciones de su organismo.

La IA es el campo de la ciencia y la ingeniería que se centra en el estudio de cómo incluir en el computador el comportamiento inteligente y la creación de artefactos, más bien “cientefactos”, que muestran tal comportamiento inteligente (60). Proporciona a las máquinas la capacidad de “razonar”, “aprender”, etc. y las utiliza en situaciones complejas donde se tiene un conjunto completo de datos, informaciones y conocimientos, con un rápido procesamiento de los mismos y una eficiente ayuda en la toma de decisiones. Se consideran datos a los valores que pueden tomar las distintas variables, informaciones cuando los datos adquieren significado y conocimiento cuando datos e informaciones adquieren una característica de utilidad.

El aprendizaje automático (ML) es considerado una disciplina científica dentro de la IA que trata de que sistemas artificiales puedan poseer la capacidad de aprendizaje, sobre todo “socrático”; esto es, a partir de ejemplos y no de instrucciones. Arthur Samuel en 1959, lo explicó como “el campo de estudio que da a las computadoras la capacidad de aprender sin que se haya programado explícitamente”(61) y ya antes, en 1943, McCulloch y Pitts propusieron el primer elemento formal o neurona formal (62).

En la base del aprendizaje automático está por un lado la representación de datos introducidos dando lugar a un conjunto principal, y por otro, la generalización, que consiste en que el sistema se ejecute adecuadamente frente a la introducción de nuevos datos.

Para ayudar en el diagnóstico clínico, ML plantea perfeccionar las técnicas ya existentes eliminando los sesgos y añadiendo un gran conocimiento del que se dispone de manera metódica y organizada.

Homer Warner en 1961 elaboró en Utah, a partir de 50 atributos clínicos, un método automatizado para el diagnóstico de 35 cardiopatías congénitas, obteniendo un resultado superior al de los cardiólogos de referencia (63). La Universidad de Stanford desarrolló MYCIN en la década de los 70, un sistema de ayuda diagnóstica en infecciones y alteraciones de coagulación. Si bien por cuestiones legales no se ejecutó, revelaba que el 65% de los tratamientos indicados eran correctos, siendo superior a los resultados de cinco profesionales expertos en infectología (entre 42,5 y 62,5%) (64).

En los años 70-80, se desarrollaron CADUCEUS/INTERNIST-1/Quick Medical Reference e Iliad, equipos educativos electrónicos que daban buenos resultados diagnosticando enfermedades en medicina interna (65). Sus principales desventajas fueron una complicada interfaz y una duración de consulta excesiva, pero han servido como antecedente evolucionado hacia nuevas iniciativas y experimentos.

Indudablemente, con el paso del tiempo se amplían las funciones destinadas al aprendizaje automático, encontrándose en continua evolución. En Estados Unidos, la Fundación Nacional para la Ciencia notificó, en el año 2012, “*Big Data*”, colaborando con los Institutos Nacionales de Salud e incluyendo capital para investigación, aprendizaje y decisiones colaborativas. La tendencia en el ámbito médico también se ha ido incrementando, así, en la Conferencia Internacional de Minería de Datos de 2012 en Siam, han prevalecido de manera considerable.

---

### 1.3.1. TÉCNICAS DE INTELIGENCIA ARTIFICIAL Y APLICABILIDAD EN BIOMEDICINA

Desde los comienzos de la IA, se orientó su evolución, entre otras áreas de aplicación, a la posibilidad de mejorar la atención en el cuidado de la salud de las personas. Las primeras aproximaciones que han tenido un cierto éxito se alcanzan en los años 70-80 del siglo XX. En España, se comenzaron a tener los primeros resultados en grupos académicos pioneros, como el GIB (Grupo de Informática Biomédica) de la FIM (Facultad de Informática de Madrid), de la Universidad Politécnica de Madrid o grupos y reuniones científicas financiadas por las más potentes empresas informáticas de la época, como IBM o Hewlett Packard.

La interrelación IA y salud evoluciona prósperamente hasta alcanzar en los últimos 15 años desarrollos específicos en Diagnóstico, Pronóstico (66), Monitorización (67), Control (68) y aspectos multidisciplinarios (69). El paradigma actual de la utilización de la IA en el ámbito de la salud es el Watson-health de IBM (<https://www.ibm.com/watson-health>). Se trata de un producto diseñado para incrementar la experiencia de los clínicos y mejorar la asistencia médica, basándose en el procesamiento inteligente de cantidades ingentes de información y con la capacidad de adquirir conocimiento; esto es, aprender, en base a la experiencia.

Sus fundamentos son:

- El procesamiento en la nube, para asegurar la calidad, el contexto, la relevancia y el control de cómo son usados los datos.

- El análisis de datos, informaciones y conocimientos para poner de relevancia lo importante en cada caso.
- El uso de técnicas y procedimientos de IA para ir mejorando las respuestas, aprendiendo de cada caso que maneja el sistema.
- La tecnología “Block-chain” que aporta los procesos de seguridad necesarios, referentes a intromisiones y, o, pérdida de datos, informaciones y conocimientos.

*En igualdad de condiciones, la explicación más sencilla suele ser la más probable.* En el contexto del aprendizaje automático, este principio metodológico conocido como la “navaja de Ockham” y atribuido al fraile franciscano Guillermo de Ockham (70), significa que, si dos sistemas tienen un rendimiento similar para unas condiciones identificadas como las más destacables, como pueden ser la precisión y la estabilidad para un problema planteado, siempre se debería de elegir el sistema “más simple”, es decir, el menos complejo de implementar y, por lo tanto, el más fácil de interpretar.

Considerando que no existe un mejor sistema para todo, en esta Tesis se comprueba el rendimiento de las “máquinas de soporte vectorial”, “Random Forest”, “k-nearest neighbor (en adelante, KNN)” y “Elastic Net”, siendo algunas de las técnicas que proporcionan mejores resultados.

### *Máquinas de soporte vectorial*

Las máquinas de vectores de soporte (en adelante, SVM) son un conjunto de algoritmos de aprendizaje supervisado que examinan datos y reconocen modelos relacionados con problemas de clasificación y análisis de regresión (71).

Este método elige un conjunto de ejemplos de entrenamiento y predice, para cada *input*, cuál de dos clases representa el *output*: es un clasificador lineal binario no probabilístico. Se representan los datos como puntos en el espacio y se busca el hiperplano que separa los datos de la forma más amplia posible.

Tenorio et al. en 2011 realizaron un método de apoyo para la decisión clínica en el diagnóstico de la enfermedad celíaca, con información de 178 sujetos y utilizando cinco métodos: árboles de decisión, inferencia bayesiana, KNN, SVM y redes neuronales artificiales. La técnica con mejores resultados fue un clasificador bayesiano, con exactitud de 80%, sensibilidad de 0.78, especificidad de 0.80 y área debajo de la curva (AUC) de 0.84 al medirse con el estándar de referencia, la biopsia intestinal (72).

En el estudio nacional hecho por Roberto González et al (73), se propone el empleo de clasificadores binarios centrados en las SVM para obtener un claro diagnóstico diferencial entre enfermos de Párkinson y de temblor esencial, llegando a la conclusión

de que se logran cubrir los objetivos con tasas de acierto del 100% en dicha diferenciación y presentando excelentes ventajas que justifican los buenos resultados conseguidos.

### *Random Forest*

Los *ensembles* de dos o más aproximaciones como Random Forest, Bagging (bootstrap aggregating) y Boosting, han obtenido mejores resultados que las técnicas individuales.

Un grupo multidisciplinario en Serbia y Polonia (74) clasificó las cefaleas a través de un test clínico poniendo a prueba varias metodologías: Naïve Bayes, SVM, Random Forest, Bagging y Boosting.

La exactitud se mantuvo entre el 71 y 81%, siendo mejor con las últimas tres técnicas. A mayores, se emplearon sistemas de selección de atributos con el fin de reducir la dimensionalidad y considerar las variables más importantes. Con dichos métodos se obtuvieron mejores resultados que con los profesionales expertos, los cuales escogieron intuitivamente las características clínicas más relevantes.

Actualmente, la cohorte inglesa del estudio de S. F. Weng (75) et al., es uno de los principales trabajos en los que se ha desarrollado un modelo de predicción de muerte prematura. Dicha cohorte prospectiva se compone por 502,628 participantes de 40 a 69 años reclutados para el Biobanco del Reino Unido entre 2006 y 2010, a la que se le ha realizado un seguimiento hasta el año 2016, aportando información sobre factores demográficos, biométricos, clínicos y sobre sus estilos de vida.

A partir de los datos de mortalidad, se trabajó con dos modelos de aprendizaje automático, Random Forest y Deep learning, enfocados a estudiar y analizar datos de casos reales. Los investigadores descubrieron que sus predicciones de muerte prematura eran muy precisas, logrando mejores resultados que con la Regresión de Cox.

El aprendizaje automático mejoró significativamente la precisión de la predicción de mortalidad prematura por cualquier causa en esta población de mediana edad, en comparación con los métodos estándar. Este estudio recrea el valor del aprendizaje automático para la predicción de riesgos dentro de un diseño de estudio epidemiológico tradicional, y cómo se podría informar este enfoque para ayudar a la verificación científica.

### *k-Nearest Neighbor (kNN)*

El razonamiento basado en casos consiste en resolver conflictos nuevos a través de casos anteriores semejantes (76), los cuales una vez solucionados pasan a la biblioteca de casos útiles para cotejar, con lo que el sistema cada vez adquiere más solidez.

En Finlandia, un grupo con una amplia experiencia en bioinformática ha llevado a cabo un abordaje de inmunodeficiencias primarias asistido con 708 casos clínicos y 148 atributos, el cual planteó el desarrollo de un sistema experto con razonamiento basado en casos. Personalmente reclasificaron las inmunodeficiencias en 11 grupos a través de un consenso resuelto con cinco técnicas (77).

En los resultados se obtuvieron 469 casos correctamente clasificados, con una exactitud del 66%. Las limitaciones de los sistemas expertos es que dependen del conocimiento que posea el experto en cuestión y de su capacidad para expresar este conocimientos en forma de reglas de producción (si...entonces) u otro sistema de representación del conocimiento que permita ser computado por un sistema artificial.

### *Elastic Net*

Elastic Net (78), es una generalización de la regresión Ridge (79) y de Lasso (80), y un importante método, para la clasificación y la regresión. Es una técnica de regularización para ajustar un modelo lineal generalizado, debido a que crea un modelo predictivo multivariado, que es susceptible de validación cruzada y puede asimilar fácilmente características continuas y categóricas. Además, puede realizar una selección de características, lo que significa que el modelo resultante puede incluir tan pocas características como sea posible, sin reducir el rendimiento del clasificador.

Elastic Net se ha utilizado en numerosas y diversas aplicaciones, incluida la identificación de marcadores genómicos de sensibilidad a fármacos (81), el desarrollo de un predictor de la edad (82) y en la determinación de factores de riesgo para el consumo excesivo de alcohol (83).

## 1.4. JUSTIFICACIÓN DEL ESTUDIO

Actualmente, existe una tendencia al alza por el estudio de graves patologías fuertemente vinculadas a factores nutricionales, siendo numerosas de naturaleza crónica y que constituyen graves inconvenientes de salud en los países desarrollados, tanto a nivel individual por su comorbilidad, como colectivo por sobrecargar el sistema asistencial de salud dañando seriamente su calidad asistencial por consumo de recursos y su sostenibilidad. Además, también supone un elevado porcentaje de bajas laborales lo que incide directamente en una merma de la competitividad del sector productivo. La valoración del estado nutricional es un aspecto muy importante en la localización de grupos de riesgo de deficiencias y excesos dietéticos que pueden ser desencadenantes o factores de empeoramiento de muchas de estas enfermedades.

Por todo lo expuesto anteriormente, es necesario realizar un análisis de los factores que inciden en el objetivo de alcanzar un mejor estado nutricional de la población.

La irrupción y el éxito alcanzado por las técnicas y procedimientos de IA para analizar datos complejos aconseja tener en cuenta su uso en el problema aquí abordado, ya que se trata de un problema complejo en el que dichas técnicas podrían aportar nuevo conocimiento clínico que, de otro modo, no parece posible alcanzar.

Así, se justifica la utilización de las técnicas y procedimientos de IA para analizar los datos, informaciones y conocimientos obtenidos en el presente estudio nutricional y de adherencia a la dieta mediterránea, y comparar los resultados con los obtenidos mediante el uso de técnicas y procedimientos estadísticos clásicos.

## 2. HIPÓTESIS Y OBJETIVOS



La detección temprana del nivel de salud y bienestar de un individuo o población es fundamental, tanto como medida de prevención primaria como secundaria. El concepto de evaluación del estado nutricional está ligado a la salud poblacional de tal manera que su estudio se manifiesta como una necesidad primordial en nuestro actual sistema asistencial.

Además, las propiedades cardiosaludables, entre otras, que se le atribuyen a la dieta mediterránea la convierten en el paradigma mundial de la dieta sana, por lo que conocer el grado de adherencia entre nuestra población se convierte en esencial.

Es importante la identificación de grupos de riesgo de deficiencias y excesos dietéticos que pueden ser desencadenantes o favorecedores de un mal pronóstico de múltiples enfermedades.

El uso de técnicas de inteligencia artificial puede resultar ventajosa, en comparación con las técnicas estadística clásicas, para la identificación de subgrupos poblacionales en riesgo de malnutrición.

Por todo ello, la hipótesis de trabajo que se plantea es la siguiente:

Los diferentes algoritmos de “Machine Learning”, en comparación con las técnicas estadísticas clásicas, permiten identificar, con un mayor valor predictivo, las características que presentan una mayor influencia en el estado nutricional y en el grado de adherencia a la dieta mediterránea en población adulta.

Para comprobar si dicha hipótesis es correcta, se ha planteado como objetivo primario, llevar a cabo un trabajo de investigación de campo para determinar la prevalencia de riesgo de malnutrición en población general adulta, su adherencia a la dieta mediterránea. Para ello, se realiza un estudio comparativo entre técnicas estadísticas clásicas y técnicas de inteligencia artificial para establecer aquellas variables más determinantes del estado nutricional y el grado de adherencia en esta población. Con dicha comparación se comprobará si ambas aproximaciones dan relevancia a las mismas variables, o si se descubren nuevas relaciones significativas entre las variables.

Como objetivo secundario, se espera que los diferentes sistemas de “Machine Learning” aporten nuevo y mayor conocimiento, nuevas perspectivas, y descubran las características que tengan un mayor valor predictivo del estado nutricional y del grado de adherencia a la dieta mediterránea en los sujetos a estudio.



### 3. METODOLOGÍA, MATERIAL Y FASES



### 3.1 METODOLOGÍA

#### 3.1.1. DISEÑO DEL ESTUDIO

La presente Tesis Doctoral desarrolla un estudio observacional transversal de prevalencia en individuos de 40 y más años. Este proyecto se llevó a cabo en el centro de salud del municipio de Cambre (Figura 1). El municipio de Cambre está localizado al noroeste de Galicia, en la provincia de A Coruña, a 12 Km en dirección sureste, de la capital de la provincia. La superficie total del municipio es de 41 km<sup>2</sup>, con una población según el Instituto Gallego de Estadística (IGE) de 23.879 habitantes datados en el año de realización de los trabajos de campo del presente estudio.

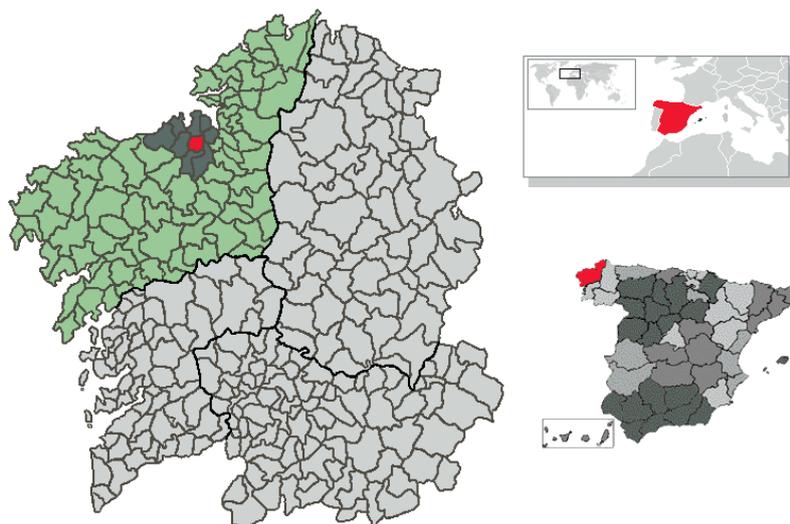


Figura 2. Localización del municipio de Cambre

Imagen tomada de *Wikimedia Commons*.

URL: [https://commons.wikimedia.org/wiki/File:Location\\_of\\_Cambre.png?uselang=es](https://commons.wikimedia.org/wiki/File:Location_of_Cambre.png?uselang=es)

La recogida de datos se realizó desde Noviembre del 2012 a Diciembre del 2013 por una enfermera y dos médicos de atención primaria previamente entrenados para la tarea a llevar a cabo.

Para recoger la información necesaria, se contactó en principio por correo postal con las personas aleatoriamente seleccionadas que cumplían los criterios de inclusión del

---

estudio, enviándoles una carta de invitación y un tríptico informativo (Anexos 1-2) al domicilio, donde se explicaba en qué consistía dicho estudio, los objetivos, los riesgos e inconvenientes. A los 3-5 días, telefónicamente, se concertó una cita en el centro de salud donde se realizó la exploración y recogida de la información. Previo consentimiento informado (Anexo 3), se realizó la entrevista personal donde se incluía la cumplimentación de cuestionarios sobre la salud de cada participante y su exploración física.

---

### 3.1.2 CRITERIOS DE INCLUSIÓN Y EXCLUSIÓN

Se han incluido a los individuos de 40 y más años que, tras una selección aleatoria, dieron su consentimiento a participar. Se excluyeron a los sujetos que no consideraron oportuno firmar el consentimiento informado y, o, que no podían acudir al centro de salud por múltiples motivos: jornada laboral, incapacidad, falta de medios de desplazamiento; y las personas que padecían un deterioro cognitivo que hacía inviable su participación en el proyecto.

---

### 3.1.3 JUSTIFICACIÓN DEL TAMAÑO MUESTRAL

Teniendo en cuenta la población de 40 y más años del Ayuntamiento de Cambre, en el año 2012, según datos del IGE (Tabla 4), a través de los datos de la tarjeta sanitaria y tras realizar un muestreo estratificado por grupos de edad y sexo, para una seguridad del 95% ( $\alpha=0,05$ ), una precisión  $\pm 4,28\%$ , ajustando un 10% de pérdidas y asumiendo que la probabilidad de los eventos es un  $\pm 50\%$ , se determina que será preciso incluir 503 sujetos en el estudio.

Tabla 4. *Datos poblacionales del municipio de Cambre (A Coruña) del año 2012 y datos muestrales según edad y sexo.*

GRUPOS DE EDAD	POBLACIÓN			MUESTRA		
	TOTAL	HOMBRES	MUJERES	TOTAL	HOMBRES	MUJERES
40-44	2465	1202 (26,9%)	1263(27,8%)	33	19	14
45-49	2231	1110(24,8%)	1121(24,7%)	85	52	33
50-54	1763	857(19,2%)	906(19,9%)	54	32	22
55-59	1383	702(15,7%)	681(14,9%)	33	18	15
60-64	1170	598(13,4%)	572(12,6%)	48	26	22
<b>Total (40-64 años)</b>	9012	4469	4543	253	147	106
65-69	1027	497(33,0%)	530(27,5%)	94	57	37
70-74	688	337(22,4%)	351(18,2%)	77	46	31
75-79	777	326(21,6%)	451(23,4%)	46	28	18
80-84	511	198(13,1%)	313(16,2%)	24	12	12
85 y más	431	148(9,8%)	283(14,7%)	9	7	2
<b>Total (65 y más años)</b>	3434	1506	1928	250	150	100

### 3.1.4 MEDICIONES

#### 3.1.4.1. Variables de identificación del paciente

Después de explicar los objetivos del estudio y firmar el consentimiento informado (Anexo 3), se adjudicó a cada paciente un código de identificación con la finalidad de asegurar la confidencialidad de la información en la totalidad del proyecto.

#### 3.1.4.2. Variables sociodemográficas

Las variables sociodemográficas que se consideró oportuno incluir para el presente estudio son (Anexos 4-5):

- Edad (años)
- Sexo
- Nivel de estudios: Sin estudios, 1º grado incompleto, 1º grado (EGB), 2º grado, universitario.
- Estado civil: Soltero, casado/pareja, viudo, divorciado, separado.
- Núcleo de convivencia (vive sólo o no)

### 3.1.4.3. Comorbilidad

Posteriormente, se recurrió a la información de la historia clínica, dónde se comprobó y completó la información con la obtenida a través de la entrevista al paciente el día del estudio, evaluando las comorbilidades del Score de Charlson (84), además de la tensión arterial y el tabaquismo.

El índice de comorbilidad de Charlson (Anexo 6), creado originalmente en 1987 (84) y adaptado posteriormente en el 1994 (85), es un sistema de evaluación de las comorbilidades del sujeto y de la esperanza de vida a los diez años independientemente de la edad en que se valora. Se ha empleado para muchos fines, entre ellos para el cálculo de costos procedentes del padecimiento de patologías crónicas en Atención Primaria (86).

A mayores de la edad, existen 19 ítems, los cuales se ha cotejado que afectan a la esperanza de vida del individuo. Cada uno de ellos se puntúa de 1, 2, 3 o 6 en función del riesgo de fallecer asociado. Posteriormente, se suman las puntuaciones y el resultado pronostica la mortalidad. Esto se presenta una vez calculados los totales en 3 categorías: ausencia de comorbilidad (0-1 puntos), baja comorbilidad (2 puntos) y alta comorbilidad ( $\geq 3$  puntos) (ver Tabla 5).

Tabla 5. *Grado de comorbilidad según el Score de Charlson*

<b>GRADO DE COMORBILIDAD</b>	<b>PUNTUACIÓN</b>
Ausencia de comorbilidad	0-1 puntos
Comorbilidad baja	2 puntos
Comorbilidad alta	$\geq 3$ puntos

En cuanto al registro de la tensión arterial, tras al menos 5 minutos de reposo, se determinaron dos tomas de tensión arterial sistólica y diastólica, una al comienzo del estudio y otra al finalizar el mismo, intentando evitar el fenómeno “hipertensión de la bata blanca” que se manifiesta cuando la presión del paciente se eleva en la consulta, pero al salir del centro sanitario desciende hasta llegar a normalizarse, distorsionando la toma de la tensión arterial.

Se realizó en la consulta del centro asistencial con el sujeto sentado en una silla con la espalda apoyada, las piernas descruzadas, los pies en el suelo y el brazo apoyado manteniendo el antebrazo a la altura del corazón. Se ha empleado un esfigmomanómetro Riester validado para uso clínico.

Posteriormente, se recogió información autodeclarada acerca del tabaquismo con las siguientes categorías: si el sujeto no fuma, si es exfumador (el paciente dejó de fumar hace más de 12 meses antes de entrar en el estudio) o si fuma actualmente, valorando el grado de dependencia mediante el Test de Fageström (87) (Anexo 7). Fageström lleva varias décadas elaborando escalas para evaluar el nivel de dependencia de la población general fumadora a la nicotina, la sustancia psicoactiva por excelencia que contiene el tabaco, pretendiendo diferenciar a los fumadores duros de los blandos. La versión más actualizada es el test de Fageström de dependencia a la nicotina. Ha sido validado con medidas fisiológicas del contenido de nicotina en sangre, permitiendo que sea un test utilizado para medir el grado de dependencia fisiológica (88).

El test de Fageström está compuesto por 6 ítems entre dos y cuatro opciones de respuesta. La puntuación varía entre 0 y 10 puntos. Una puntuación de 6 o más indica que el sujeto tiene una elevada dependencia fisiológica a la nicotina, mientras que una cifra de 5 o menos, es señal de dependencia media o baja. En la Tabla 6, se muestra el modo de puntuación, para conocer si el fumador tiene dependencia baja, media o alta de la nicotina.

Tabla 6. *Dependencia a la nicotina según el Test de Fageström*

<b>GRADO DE DEPENDENCIA</b>	<b>PUNTUACIÓN</b>
Fumador ligero	2-3 puntos
Fumador normal	4-5 puntos
Fumador dependiente	6-7 puntos

Los puntos de corte son 4 y 7. Menor de 4 indica una dependencia baja. Entre 4 y 7 puntos significa una dependencia moderada. Más de 7 puntos se interpreta como una dependencia alta (87,89–91).

#### 3.1.4.4. Variables antropométricas

- **Peso (kg).** La medición del peso corporal se llevó a cabo con una balanza con tallímetro modelo MB 201T plus Asimed manual, con una capacidad máxima de 150 kg y una precisión de 100 gramos, según especificaciones del equipo. Para ello, se solicitó a los pacientes que se descalzasen, retirasen la ropa de abrigo y vaciasen sus bolsillos.
- **Talla (cm).** Se midió al mismo tiempo que el peso, con el mismo instrumento, con el sujeto descalzo y con la espalda en contacto con el tallímetro.
- **Índice de masa corporal (IMC)** o Índice de Quetelet. Se calcula a partir del peso y la talla del paciente según la fórmula (EQ.1):

$$(EQ. 1) \quad IMC = \frac{\text{Peso (Kg)}}{\text{Talla(m)}^2}$$

El sobrepeso y la obesidad se clasificaron según los criterios de la Organización Mundial de la Salud (OMS), siendo aplicable tanto a hombres como a mujeres en edad adulta. El rango para la normalidad se limita a valores de IMC entre 18,5 – 24,9 kg/m<sup>2</sup>, considerando sobrepeso valores entre 25-29,9 kg/m<sup>2</sup>; y siendo el punto de corte de la obesidad 30 kg/m<sup>2</sup>. A partir de un IMC > 30 kg/m<sup>2</sup> se entran en los diferentes tipos de severidad de obesidad. (Tabla 7) (92,93).

Tabla 7. *Clasificación de la obesidad según la OMS*

CATEGORÍA	VALORES LÍMITE DEL IMC (KG/M <sup>2</sup> )
Peso insuficiente	<18,5
Normopeso	18,5-24,9
Sobrepeso	25-29,9
Obesidad de tipo I	30-34,9
Obesidad de tipo II	35-39,9
Obesidad de tipo III (extrema)	≥ 40

- **Perímetro de cintura** (cm). El perímetro abdominal se calcula colocando al individuo en bipedestación, con los pies juntos, el abdomen relajado y los brazos colgando junto a los costados. A la altura del ombligo entre los márgenes costales inferiores y las crestas ilíacas se rodea con cinta métrica inelástica y sin presionar.

Ha sido necesario indicar valores umbral del perímetro de cintura ya que el grado de obesidad central a partir del cual aumenta el riesgo cardiovascular y metabólico varía en función de la etnia. Así, en la Tabla 8 se indican los puntos de corte admitidos por la OMS para la raza europea (94,95).

Tabla 8. *Puntos de corte del perímetro de cintura (OMS)*

Circunferencia cintura	Hombres	Mujeres
Normal	<94cm	<80cm
Riesgo elevado	94-101,9cm	80-87,9 cm
Riesgo muy elevado	≥102 cm	≥88cm

- **Perímetro de cadera** (cm). Es el perímetro medido en centímetros que pasa por la zona glútea más prominente. Se realizó con una cinta métrica inelástica con el sujeto erguido, relajado y con las rodillas juntas.
- **Índice cintura/cadera (ICC)**: A través de este índice se conoce la distribución de la grasa corporal total. En función de la distribución se clasifican dos tipos de obesidad: androide (intrabdominal o visceral) y ginecoide (extrabdominal o subcutánea). Esta medida es complementaria al IMC, ya que éste no identifica si el sobrepeso deriva de una hipertrofia muscular fisiológica o de un incremento de grasa corporal patológica.

Para calcular el ICC se divide el perímetro de la cintura por el perímetro de la cadera (en centímetros). La OMS propone unos umbrales de normalidad para este índice de 0,8 en mujeres y 1 en varones. Mediciones superiores determinarían obesidad abdominovisceral (*central*), lo cual se asocia a un riesgo cardiovascular elevado y a una mayor probabilidad de padecer enfermedades como diabetes mellitus o hipertensión arterial (94).

- **Medición de los pliegues cutáneos** (mm): La masa grasa está formada especialmente por tejido adiposo subcutáneo y perivisceral. En el adulto sano, tiene valores entre 10 y 20% en el hombre y de 15 a 30% en la mujer (96), constituyendo el 50% de la grasa corporal.

La valoración de los pliegues cutáneos se realiza en unas zonas determinadas, tal como se expone a continuación. Para esta medición, se ha empleado el plicómetro digital modelo Trimeter que determina el grosor de una doble capa de piel y tejido adiposo subyacente (sin incluir el músculo). Este plicómetro tiene un alcance de hasta 70 mm y ejerce una presión constante de 10 gr/mm<sup>2</sup> con una precisión de 0,1 mm según las especificaciones del equipo.

Los pliegues elegidos para medir la grasa subcutánea fueron:

- **Pliegue cutáneo tricipital (PCT)**: a través de este pliegue se obtiene información del contenido graso subcutáneo y el estado de las reservas corporales de energía. La medida se realizó con el individuo en bipedestación, brazos relajados, en el punto medio posterior del miembro superior no dominante, entre el acromion y el olécranon.
- **Pliegue cutáneo bicipital**: se mide en el punto medio de la cara anterior del brazo, a la misma altura del PCT y con la misma técnica, a mitad de distancia entre olécranon y el acromion, con el codo flexionado a 90°.
- **Pliegue cutáneo subescapular**: este pliegue mide la obesidad troncular. Primeramente, se localiza el ángulo inferior de la escápula. Posteriormente, se coloca el dedo índice justo debajo y con el pulgar se hace pinza de manera oblicua, así la medida del pliegue será diagonal sobre un ángulo de 45°.
- **Pliegue cutáneo suprailíaco**: para medir este pliegue se determina una zona sobre la parte superior de la cadera, en la línea axilar media y se mide diagonalmente sobre la zona más alta del arco.
- **Perímetro de pantorrilla (cm)**: es la circunferencia de la zona más ancha que existe entre el tobillo y la rodilla, tomada con una cinta métrica. Existe una buena correlación con la masa libre de grasa y la fuerza muscular (97–99).
- **Perímetro del brazo (PB)**: junto con la medida del PCT se puede hallar, mediante fórmulas, el compartimento muscular, graso y proteico. Se realizó con cinta antropométrica inextensible y se midió en centímetros. La medida fue tomada con la cinta antropométrica, en el punto medio del brazo, sin comprimir, en el mismo sitio donde se midió el PCT.
- **Perímetro muscular del brazo (PMB)**: a través de este perímetro se conoce la masa muscular esquelética de los sujetos (compartimento proteico). El cálculo se realiza con las medidas del PCT y PM aplicando la ecuación EQ. 2:

$$\text{(EQ. 2)} \quad \text{PMB} = \text{PB} - (\text{PCT} \times \pi)$$

- **Área muscular del brazo (AMB):** implica el compartimento muscular y calcula la reserva proteica basándose en las medidas de PB y PTC. Se calcula mediante la ecuación EQ.3:

$$\text{(EQ. 3)} \quad \text{AMB} = [\text{PB} - (\text{PCT} \times \pi)]^2 / 4\pi.$$

- **Área grasa del brazo (AGB):** determina el compartimento grasa y la reserva energética a través del área total braquial (ATB) y el área muscular del brazo (AMB). Se calcula mediante la ecuación EQ. 4:

$$\text{(EQ. 4)} \quad \text{AGB} = \text{ATB} - \text{AMB}$$

En donde  $\text{ATB} = \text{PB}^2 / 4\pi$

- **Índice Adiposo Muscular (IAM):** este índice, planteado por Alastrué (93) determina el estado nutricional a partir de las áreas adiposa y muscular del brazo. Fundamentalmente atribuido a la valoración de la obesidad, ha tenido otras finalidades como factor pronóstico en la progresión de patologías crónicas. Se obtiene mediante la ecuación EQ. 5:

$$\text{(EQ 5.)} \quad \text{IAM} = \text{AGB} / \text{AMB}$$

- **Métodos Bioeléctricos (100–102):** Otra técnica para evaluar la grasa corporal es la bioimpedanciometría eléctrica (BIA). Está basada en la variación de resistencia que existe entre los distintos compartimentos corporales al paso de una corriente alterna de baja intensidad y alta frecuencia. Esta corriente se transmite a través del agua y de los electrolitos, por lo que variará en función del contenido de grasa corporal.

En la actualidad, hay una gran variedad de analizadores de impedancia; en nuestro estudio se empleó el bioimpedanciómetro marca Beurer y modelo BG55, que posee una capacidad máxima de 150 kg y una precisión de 0,1% para grasa corporal, agua corporal y porcentaje muscular y 100 g para peso corporal, según información facilitada por el propio fabricante (103).

Con este método, se obtienen valores de peso (kg), masa adiposa (%), masa líquida (%), masa muscular (%), metabolismo basal calórico (kcal), masa ósea (kg) y metabolismo de actividad (kcal).

### 3.1.4.5 Valoración del estado nutricional

Coincidiendo con la entrevista, se les solicitó a los participantes que cubriesen el cuestionario “Mini Nutritional Assessment” (en adelante, MNA) (104,105) (Anexo 8), siendo la herramienta de cribado empleada con más frecuencia para determinar el estado nutricional, ayudando a identificar a pacientes malnutridos o en riesgo de malnutrición.

El MNA comprende 18 “ítems” que tratan los siguientes apartados: *a)* medidas antropométricas (peso, talla, y pérdida de peso); *b)* informe dietético (8 cuestiones referentes al número de tomas, ingesta de alimentos y líquidos, y autonomía en la alimentación); *c)* valoración global (6 preguntas sobre el estilo de vida, tratamientos farmacológicos y movilidad), y *d)* una evaluación subjetiva (autopercepción de salud y nutrición). En la Tabla 9, se observa la clasificación del estado nutricional en función de los puntos obtenidos. Este cuestionario obtiene una sensibilidad del 96%, una especificidad del 98% y un valor predictivo del 97% (106,107).

Tabla 9. *Clasificación del estado nutricional según el cuestionario Mini Nutritional Assessment (MNA)*

ESTADO NUTRICIONAL	PUNTUACIÓN
Estado nutricional normal	≥ 24 puntos
Riesgo nutricional	17-23,5 puntos
Malnutrición	< 17 puntos

### 3.1.4.6 Test de adherencia a la dieta mediterránea

El modelo dietético de los países mediterráneos se fundamenta en una alta ingesta de verduras y hortalizas, legumbres, frutas, frutos secos y cereales y aceite de oliva; un consumo moderado de pescados, huevos y productos lácteos y un menor consumo de carnes y grasas animales (108–111).

Como ya se ha explicado, la dieta mediterránea compone un patrón de dieta saludable, garantizando un aporte calórico y nutritivo en dosis y proporciones adecuadas (52,112–114); contribuyendo a la prevención de patologías cardiovasculares, hipertensión, diabetes, cáncer, y, por lo tanto, a una esperanza de vida superior (111,115–120).

Para valorar la adherencia a la dieta mediterránea se empleó un cuestionario validado con 14 preguntas (Anexo 9), muy rápidas de responder, que permite saber si la dieta que se realiza de manera habitual puede considerarse que sigue los parámetros de la dieta mediterránea (121). Cada pregunta contestada afirmativamente suma un punto. Se considera que la persona sigue correctamente la dieta mediterránea cuando su puntuación es igual o superior a 9 puntos.

---

#### 3.1.4.7 Autopercepción de la imagen corporal

Se determinó el peso subjetivo siguiendo el modelo formulado por Marco et al. (122), para ello la persona debe elegir una de las siguientes alternativas: “*considero que estoy en mi peso;*” “*considero que tengo kilos de más*” o “*considero que tengo kilos de menos*”. Con ello, se clasificó a la población en tres grupos: “*peso-subjetivo-justo*” aquellos que afirman estar en su peso, “*kilos de más-subjetivo*” para los que creen que exceden de peso y “*kilos de menos-subjetivo*” para los que opinan que pesan de menos.

En cuanto a la medición de la imagen corporal, se les facilitó a los participantes en el estudio un cuestionario validado en España por Corral et al. (123) con cuestiones de dos de las once subescalas del “Eating Disorders Inventory-2” (124) (Anexo 10): *Insatisfacción corporal* (EDI-IC) y *Obsesión por la delgadez* (EDI-OD), por ser éstas las que evalúan específicamente la imagen corporal a través de aspectos relacionados directamente con alteraciones perceptuales.

Dichos ítems aparecen entremezclados en un único cuestionario, evitando que los individuos reconozcan el aspecto que se valora. La contestación de los ítems y la manera de corregirlos es la planteada en el manual.

- La subescala de la Insatisfacción Corporal (EDI-IC): La insatisfacción corporal puede ser un motivo de comienzo y permanencia de hábitos orientados a controlar el peso en individuos con alteraciones alimentarias. Esta subescala mide la insatisfacción generada hacia la constitución general corporal o hacia aquellos puntos que más perturban a los que sufren dichos trastornos.

Los “ítems” que componen la subescala de Insatisfacción Corporal del EDI-2 son:

- 1.- Creo que mi estómago es demasiado grande.
- 2.- Pienso que mis muslos son demasiado gruesos.
- 3.- Creo que mi estómago tiene el tamaño adecuado.
- 4.- Me siento satisfecho con mi figura.
- 5.- Me gusta la forma de mi trasero.

- 6.- Creo que mis caderas son demasiado anchas.
- 7.- Creo que el tamaño de mis muslos es adecuado.
- 8.- Creo que mi trasero es demasiado grande.
- 9.- Creo que mis caderas tienen el tamaño adecuado.

- La subescala de la Obsesión por la Delgadez (EDI-OD): Bruch (125) definió esta subescala como un factor esencial de las alteraciones en el comportamiento alimentario hacia el peso, las dietas y el miedo a engordar.

Los “ítems” que forman la subescala de Obsesión por la Delgadez del EDI-2 son:

1. Como dulces y carbohidratos sin preocuparme.
2. Pienso en ponerme en dieta.
3. Me siento muy culpable cuando como en exceso.
4. Me aterroriza la idea de engordar.
5. Exagero o doy demasiada importancia al peso.
6. Estoy obsesionado por el deseo de ser más delgado.
7. Si aumento un kilo de peso temo que seguiré engordando.

---

### 3.1.5 ASPECTOS ÉTICO-LEGALES

- El estudio se realizó siguiendo las normas de buena práctica clínica de la Declaración de Helsinki.
- Se garantizó la confidencialidad de la información recogida según la Ley Orgánica 15/1999 de Protección de Datos de Carácter Personal y el actual Reglamento Europeo en materia de Protección de Datos.
- Se garantizaron en todo momento los requisitos éticos y legales exigidos por la normativa de investigación biomédica vigente aplicable al caso concreto. El acceso a la historia clínica se realizó de acuerdo con la legislación vigente:
  - Ley 41/2002 básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica.
  - Ley 3/2005, de modificación de la Ley 3/2001, reguladora del consentimiento informado y de la historia clínica de los pacientes.
  - Decreto 29/2009, de 5 de febrero, por el que se regula el uso y acceso a la historia clínica electrónica.

- Todos los participantes otorgaron su consentimiento informado para participar en el mismo. Este estudio cuenta con la aprobación del Comité Ético de Investigación Clínica de Galicia (códigos autorización 2008/264 y 2012/390) (Anexos 11-12)

---

### 3.1.6 ANÁLISIS DE LOS DATOS OBTENIDOS EN EL TRABAJO DE CAMPO

---

#### 3.1.6.1. Técnicas “clásicas” de análisis estadístico

El análisis estadístico se realizó con el programa SPSS 24.0 para Windows.

Se realizó un estudio descriptivo de todas las variables incluidas en el estudio. Las variables cuantitativas se expresan como media±desviación típica (D.T.) y mediana. Las variables cualitativas se expresan como valor absoluto (n) y porcentaje, con la estimación de su 95% intervalo de confianza (I.C.).

Se determinaron las variables asociadas al riesgo de malnutrición y a la adherencia a la dieta mediterránea en un análisis bivariante y multivariante.

La comparación de medias entre grupos se realizó por medio del test T de Student y ANOVA o el test de Mann-Whitney o test de Kruskal-Wallis según procediese, tras comprobación de la normalidad con el test de Kolgomorov- Smirnov. La asociación de variables cualitativas se estimó por medio del estadístico Chi-cuadrado o el test exacto de Fisher. Se calcularon los valores de odds ratio (OR) y su 95% intervalo de confianza.

En el análisis multivariante, se realizaron modelos de regresión logística ajustando por aquellas variables que resultaron estar significativamente asociadas a las variables respuesta en el análisis bivariado.

Todos los test se realizaron con un planteamiento bilateral, considerándose significativos valores de  $p < 0.05$ .

---

#### 3.1.6.2 Técnicas de inteligencia artificial (Machine Learning)

Se utilizó el entorno de programación R (126) y el paquete mlr (127).

R es uno de los sistemas de software más populares y ampliamente utilizados para estadística, extracción de datos y aprendizaje automático. Sin embargo, no define un interfaz estandarizado para, por ejemplo, un modelado predictivo supervisado.

Para cualquier experimento complejo, es necesario escribir un código extenso, tedioso y propenso a errores. El paquete “mlr” ofrece un lenguaje de dominio limpio, fácil de usar y flexible para experimentos de aprendizaje automático en R.

El paquete “mlr” proporciona un marco genérico, orientado a objetos y extensible para la clasificación, regresión, análisis de supervivencia y agrupación para el lenguaje R. Proporciona una interfaz unificada a más de 160 técnicas de modelado e incluye meta-algoritmos y técnicas de selección de modelos para mejorar y ampliar la funcionalidad con, por ejemplo, ajuste de hiperparámetros, selección de características y construcción de conjuntos. El paquete se dirige a los profesionales que desean aplicar rápidamente algoritmos de aprendizaje automático, así como a los investigadores que desean implementar y comparar sus nuevos métodos en un entorno estructurado.

## 3.2 MATERIAL Y FASES

La experimentación es el método empírico que permite obtener los resultados necesarios para poder validar o refutar hipótesis y, por ello, se utiliza en la mayoría de las investigaciones científicas e industriales. Determinar correctamente las fases de un diseño experimental es la manera más eficaz, sino la única, de garantizar resultados correctos, resolviendo así el problema propuesto.

Primero, se plantean dos fases comunes que comparten la estadística clásica y las técnicas de ML, dónde se seleccionan y preparan los datos recogidos creando una muestra correcta y significativa sobre la que trabajar: selección de datos y preprocesado de los mismos.

Posteriormente, se exponen dos fases específicas para ML: el aprendizaje y la selección del mejor modelo predictivo.

---

### 3.2.1 FASES COMUNES DEL DISEÑO EXPERIMENTAL

---

#### 3.2.1.1 Selección de datos

En primer lugar, se debe generar un *dataset*, garantizando que sea lo suficientemente representativo, teniendo en cuenta sobre todo tres dimensiones de suma importancia: confiabilidad, inferencia estadística y objetividad (128).

- La confiabilidad es la capacidad de obtener resultados estables, consistentes e iguales a sí mismos en mediciones sucesivas, en diferentes tiempos, y previsibles para el futuro. Es externa cuando investigadores independientes llegan a los mismos resultados al estudiar una realidad en tiempos o situaciones diferentes, e interna cuando diferentes observadores coinciden en los hallazgos al estudiar el mismo fenómeno.
- La inferencia estadística es el conjunto de métodos y técnicas que permiten sacar conclusiones a partir de la información empírica proporcionada por el *dataset*, y ver cuál es el comportamiento de dicha población con un riesgo de error medible en términos de probabilidad.

- La objetividad se refiere al grado de permeabilidad en cuanto a la influencia de los sesgos y tendencias del investigador/es que califican e interpretan el conjunto de datos.

Por ello, en esta Tesis Doctoral, los expertos clínicos primero deben identificar a la población de estudio y definir sus características particulares, extrayendo las variables utilizadas de forma pormenorizada y formando así la muestra que mejor represente el problema en cuestión.

---

### 3.2.1.2. Preprocesado

El preprocesamiento de datos (129) es una etapa fundamental, cuyo objetivo principal es obtener un conjunto de datos final que sea de calidad y útil para la fase de extracción de conocimiento.

Esta etapa se encarga de la limpieza de datos, su integración, su transformación y su reducción, para poder ser utilizados por las herramientas previstas.

#### 3.2.1.2.1. *Preparación de los datos*

---

En esta etapa del diseño los datos se presentan en estado crudo o puro, siendo difíciles de analizar, por lo que es esencial una revisión que verifique que no constan datos con información incompleta, valores atípicos o ruido.

Para ello, uno de los aspectos importantes es estudiar la correlación que existe entre las variables de entrada, ya que si están muy correlacionadas entre sí habrá que suprimirlas para evitar ruido en el experimento.

En esta Tesis se ha empleado la prueba T para cuantificar la correlación entre cada característica y la clase.

Tras esta preparación, valorada como el 80% del esfuerzo de un análisis (130), se considera que los datos están dispuestos para su correcto estudio.

---

### 3.2.2 FASES ESPECÍFICAS DEL DISEÑO EXPERIMENTAL DE ML

Las técnicas de aprendizaje automático aprenden de los ejemplos con las que son entrenadas, siendo de vital importancia asegurar que los algoritmos no aprendan los datos de los que disponen y que no generalicen y predigan de manera correcta futuros nuevos datos desconocidos.

---

#### 3.2.2.1. Normalización

Una vez preparados los datos, se debe realizar una normalización de los mismos. Las variables suelen aparecer en distintas escalas o tamaños, dificultando su comparación; por lo que son necesarias técnicas de estandarización que permiten hacerlas equiparables, realizando un re-escalado lineal de las variables de entrada y eliminando los *outliers* (valores atípicos).

Las nuevas variables generadas tendrán una media ( $\mu$ ) igual a cero y una desviación típica ( $\sigma$ ) igual a uno, y por lo tanto serán independientes de la unidad de medida.

---

#### 3.2.2.2 Aprendizaje

Para problemas de clasificación, se pueden aplicar las técnicas de aprendizaje máquina, identificando así si un determinado ejemplo pertenece a cierta clase o no.

Posteriormente, utilizando diferentes medidas de acierto, se procede a medir el rendimiento del algoritmo, determinando la eficiencia de las técnicas empleadas durante la fase de entrenamiento.

Se busca, además, que el modelo no se sobreentrene; es decir, que se ajuste a los datos disponibles, para asegurar así los correctos resultados empleando datos desconocidos; siendo esto posible a través de procedimientos como la validación cruzada, que se verá detalladamente en el siguiente punto.

Y, por último, es necesario el uso de técnicas para reducir la dimensionalidad, ya que cuanto mayor sea ésta más cantidad de ejemplos se precisan para el aprendizaje; y, además, introduce ruido y empeora la capacidad general del sistema.

### 3.2.2.2.1. Entrenamiento del modelo

Como se ha mencionado anteriormente, es imprescindible que en la etapa de aprendizaje no se produzca “sobreentrenamiento”, para lo que son útiles las técnicas de validación cruzada también conocidas como “cross validation”.

En la validación cruzada de K iteraciones o “K-fold cross-validation” (131) los datos de muestra se dividen en K subconjuntos. Uno de los subconjuntos se dispone como datos de prueba (validación) y el resto (K-1) como datos de entrenamiento. El procedimiento de validación cruzada es repetido durante “k iteraciones”, con cada uno de los posibles subconjuntos de datos de validación.

Para obtener un único resultado, se calcula la media aritmética de los resultados de cada iteración. Esta técnica es muy precisa ya que se calcula a partir de “K” combinaciones de datos de entrenamiento y de prueba. La selección de la cantidad de iteraciones es en función del tamaño del *dataset*, siendo la más común la validación cruzada de 10 iteraciones (“10-fold cross-validation”).

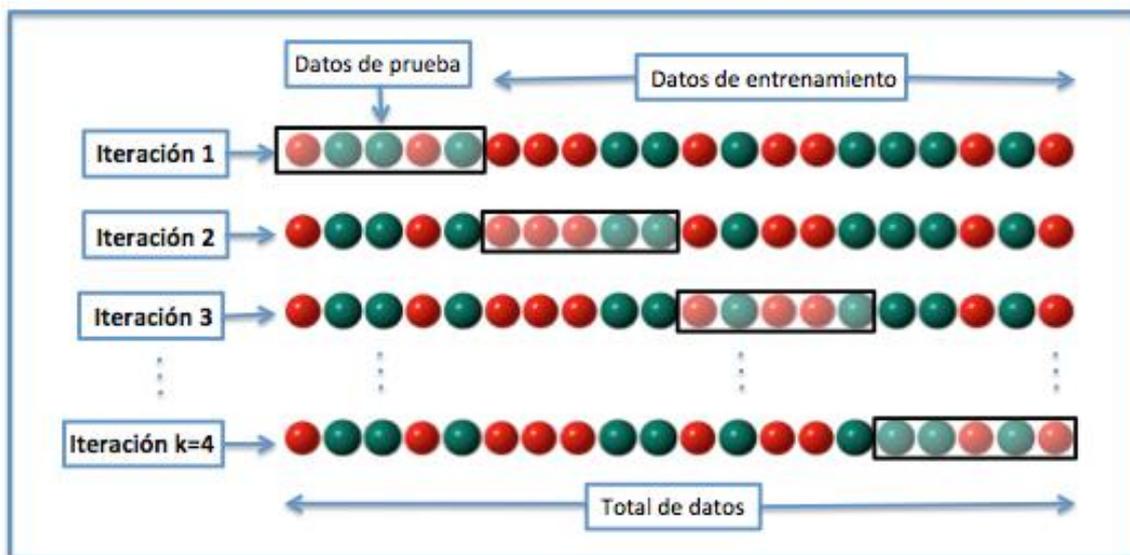


Figura 3. Validación cruzada de k iteraciones

Imagen tomada de *Wikimedia Commons*.

URL: [https://upload.wikimedia.org/wikipedia/commons/f/f2/K-fold\\_cross\\_validation.jpg](https://upload.wikimedia.org/wikipedia/commons/f/f2/K-fold_cross_validation.jpg)

Al realizar 10 iteraciones y tener un % de predicción en cada iteración en una tabla, puede representarse el resultado y obtener un promedio de la eficiencia general del modelo predictivo.

#### 3.2.2.2.2 *Reducción de la dimensionalidad*

---

En matemáticas y estadística, la *maldición de la dimensionalidad* (efecto Hughes (132)) es un problema que emerge al analizar y reorganizar datos de espacios de numerosas dimensiones.

Conforme aumenta la dimensionalidad, aumenta exponencialmente el volumen del espacio, dando lugar a que los datos disponibles se vuelvan dispersos y se dificulte la obtención de resultados con significancia estadística (133). Por lo que, para lograr un resultado sólido y fiable, el número de datos necesarios debe incrementarse exponencialmente con la dimensionalidad.

La reducción es comúnmente usada para disminuir la dimensión del conjunto de datos, desechando cualquier característica redundante o irrelevante. Este proceso de automatización es llamado selección de características, donde los principales objetivos son los siguientes: evitar el sobreajuste y mejorar el rendimiento del modelo, proporcionar modelos más rápidos y rentables y, además, obtener una visión más profunda de los procesos subyacentes que generaron los datos.

“Los árboles de decisión” y los “KNN” son clasificadores sensibles a las características irrelevantes o ruido. La técnica de Naïve Bayes es sensible a las características redundantes (134).

Hay tres aproximaciones en ML para llevar a cabo este proceso: indirecta (filtering), directa (wrapping) e intrínseca (embedding) (135–137).

- Filtrado (filtering): Selecciona las características de forma independiente del clasificador, usando un criterio de “relevancia”.
- Encapsulado (wrapping): Selecciona los subconjuntos de características en función del desempeño de un clasificador. Es costoso computacionalmente y necesita una estrategia de búsqueda para explorar de forma eficiente el espacio de subconjuntos.

- Intrínseco (embedding): Realiza la selección en el proceso de aprendizaje devolviendo un subconjunto de características y el clasificador entrenado, evaluando el coste de agregar o quitar características, pero no el reentrenamiento.

En el presente trabajo, se eligió el uso de una aproximación *filter* que, por su velocidad e independencia del clasificador, se ajusta mejor al problema que se pretende abordar (138).

Los métodos *filter* están basados habitualmente en un “ranking” de las variables predictoras, en función de un criterio de relevancia anteriormente fijado entre una variable predictora y la variable clase, a partir del cual las variables predictoras quedan ordenadas en una lista desde la más relevante a la menos relevante.

Este tipo de aproximación, se considera una selección indirecta ya que el criterio de relevancia con el que se han clasificado las variables no tiene en consideración el paradigma (Naïve Bayes, árbol de clasificación, ...).

Además, la reducción de datos puede ser alcanzada con la extracción de características, transformando el conjunto inicial en un número reducido de nuevos atributos. El “Análisis de Componentes Principales” (en adelante, PCA) es un método común usado para la reducción de datos y pertenece a las denominadas técnicas de extracción.

### 3.2.2.2.3 Estimación del error real del modelo

Una de las principales etapas de este diseño es la evaluación del rendimiento del modelo, ya que permite conocer el nivel de acierto de las predicciones del conjunto de datos mediante el modelo entrenado.

La “matriz de confusión” (Figura 4), se emplea para disponer de una visión más completa al evaluar dicho rendimiento.

		ESTIMADO POR EL MODELO	
		+	-
REALIDAD	+	<b>VERDADERO POSITIVO</b>	FALSO NEGATIVO
	-	FALSO POSITIVO	<b>VERDADERO NEGATIVO</b>

Figura 4. Matriz de confusión

Las principales *métricas de clasificación* usadas en Machine Learning son:

- Matriz de confusión o error
- Precisión
- Sensibilidad o TPR (Tasa positiva real)
- Especificidad o TNR (Tasa negativa real)
- Puntuación F1
- Área bajo la curva Característica Operativa del Receptor (AUROC)
- Pérdida logarítmica
- Cohen's Kappa

Algunos autores (139) proponen que la medida más segura es AUROC, avalada generalmente por la comunidad científica como una buena medida de acierto del rendimiento de un algoritmo.

Una curva ROC es un gráfico que muestra el rendimiento de un modelo, representando las tasas de verdaderos positivos (sensibilidad) frente a los falsos positivos (especificidad) en diferentes umbrales de clasificación.

Para reducir el umbral, se clasifican más elementos como positivos, por lo que aumentarán tanto los falsos positivos como los verdaderos positivos. En la Figura 5, se muestra una curva ROC con un valor diagnóstico perfecto.

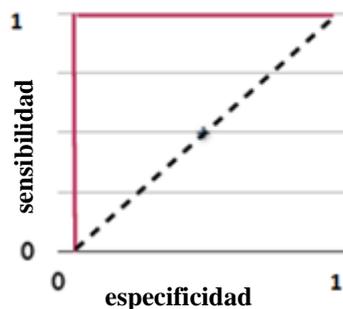


Figura 5. Curva ROC perfecta

En cuanto al AUC, es una porción del área del cuadrado unidad por lo que su valor siempre está comprendido entre 0 y 1. Sin embargo, debido a que la clasificación aleatoria produce la diagonal entre (0,0) y (1,1), lo que significa un área de 0,5 (Figura 6); ningún clasificador realista debería tener un área menor de 0,5 (139).

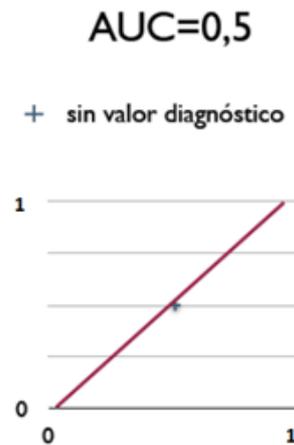


Figura 6. AUROC sin valor diagnóstico

Por un lado, las *métricas de regresión* que se han utilizado son:

- Raíz del Error cuadrático medio (RMSE)
- $R^2$
- Error cuadrático medio (MSE)

La métrica más frecuentemente empleada para los problemas de regresión es la raíz del error cuadrático medio (RMSE), que representa la raíz cuadrada de MSE (EQ 1), distancia cuadrada promedio entre el valor real y el valor pronosticado. Es una medida absoluta de ajuste que indica la proximidad entre los puntos de datos observados y los valores predichos del modelo; teniendo la utilidad de estar en las mismas unidades que la variable de respuesta.

$$(EQ. 6) \quad RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2}$$

Donde  $y_j$  es el resultado real esperado y  $\hat{y}_j$  es la predicción del modelo.

Los valores más bajos de RMSE indican un menor ajuste, por lo que es una buena medida de precisión con la que el modelo predice la respuesta.

Por otro lado, el R-cuadrado indica la bondad o aptitud del modelo, mostrando que las variables independientes seleccionadas explican la variabilidad en sus variables dependientes.  $R^2$  se define como:

$$(EQ. 7) \quad R^2 = 1 - \frac{\sum(y_j - \hat{y}_j)^2}{\sum(y_j - \bar{y}_j)^2}$$

La escala de  $R^2$  va de  $-\infty$  a 1. La mejora en el modelo de regresión da como resultado aumentos proporcionales en  $R^2$ .

---

### 3.2.2.3. Selección del mejor modelo

Para la selección del mejor modelo, con el fin de garantizar que el rendimiento de una determinada técnica de ML es o no estadísticamente relevante, se ha realizado un contraste de hipótesis.

Mediante esta teoría, se trata el problema estadístico contemplando una hipótesis nula  $H_0$  (formulada con la intención de rechazarla, siendo la que se somete a comprobación) y una hipótesis alternativa  $H_1$  (diseñada con el objetivo de aceptarla, denominándose hipótesis de investigación), y se pretende averiguar cuál de las dos es la verdadera, tras aplicar el problema estadístico a un cierto número de experimentos.

Los errores que se pueden cometer durante este proceso son de dos tipos (Tabla 10):

- Error tipo I, que se comete cuando se rechaza la  $H_0$ , siendo ésta cierta. En este caso el investigador concluye afirmando que la técnica es relevante cuando en realidad no lo es.

La probabilidad de cometer este error es el riesgo “ $\alpha$ ”, o lo que es lo mismo, la probabilidad “ $p$ ” de rechazar  $H_0$  cuando es cierta. Este riesgo se fija antes de iniciar el proceso, siendo generalmente  $p < 0,05$ .

- Error tipo II, se comete cuando no se rechaza la  $H_0$ , siendo ésta falsa. En este caso el investigador concluye que la técnica no es significativa cuando en realidad sí lo es.

La probabilidad de cometer este error es el riesgo  $\beta$  (probabilidad de no rechazar  $H_0$  cuando  $H_1$  es cierta). Su complemento se conoce como “potencia del contraste” o “poder estadístico”.

Potencia (P) =  $1 - \beta$ , que es la capacidad del estudio para detectar diferencias cuando existen en realidad.

Ambos errores están relacionados en sentido inverso, ya que al descender uno aumenta el otro.

Tabla 10. *Tipos de error en el contraste de hipótesis*

DECISIÓN	HIPÓTESIS VERDADERA	
	$H_0$	$H_1$
Aceptar $H_0$	Decisión correcta $1 - \alpha$	<b>Error tipo II</b>
Rechazar $H_0$	<b>Error tipo I</b>	Decisión correcta $1 - \beta$

Una vez definida  $H_0$ , se selecciona el test estadístico que permita evaluar su validez. La elección de esta prueba se realiza en función del número de variables a comparar, dividiéndose en: test paramétricos y test no paramétricos (140).

Para poder utilizar un test paramétrico se considera que, o se ha obtenido una muestra aleatoria que cumple las características de una distribución normal, o se ha comprobado la independencia (satisfecha con la separación mediante validación cruzada de los datos), la normalidad (utilizando el test de Shapiro Wilk (141)) y con la heterocedasticidad de los datos (a través del test de Bartlett (142)).

Si los resultados cumplen con las 3 condiciones citadas anteriormente, se aplicará una prueba paramétrica (ANOVA), de lo contrario, la versión no paramétrica (Friedman (143))

---

### 3.2.2.4 Técnicas de Machine Learning

#### 3.2.2.4.1 Máquinas de soporte vectorial

---

Las máquinas de soporte vectorial (144–146) (en adelante, SVM, por sus siglas en inglés, "Support Vector Machine"), fueron desarrolladas en los años 90 por Vapnik y cols. (147); donde adquirieron fama tras obtener mejores resultados que las redes de neuronas artificiales clásicas en el análisis de letra manuscrita ("neocognitron", propuesto por Kunihiko Fukushima en 1979 (148), y que actualmente está planteando las "Deep Convolutional Neural Networks"), usando píxeles como entrada (149).

Aunque originariamente fueron creadas para solventar problemas de clasificación binaria, en la actualidad se emplean en problemas de regresión, multclasificación y clustering; siendo también utilizadas con éxito en campos como la visión artificial, categorización de texto e hipertexto, procesamiento de lenguaje natural, análisis de series temporales, ...

Las principales ventajas que tiene SVM son:

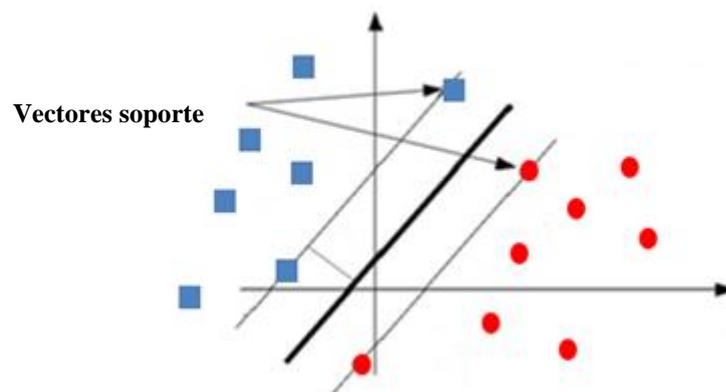
- Una capacidad muy elevada de generalización por la minimización del riesgo estructural.
- Escasos parámetros para ajustar; el modelo únicamente depende de los datos con mayor información (vectores de soporte).
- La estimación de los parámetros se realiza a través de la optimización de una función de costo convexa, lo cual evita la existencia de un mínimo local.
- La solución de SVM es *sparse*, siendo la mayoría de las variables cero en la solución de SVM, por lo que el modelo final es una mezcla de un número muy pequeño de vectores de soporte.
- La complejidad del clasificador depende del número de vectores que fijan la frontera y no de la dimensión del espacio.

Además, las SVM son clasificadores lineales que permiten encontrar hiperplanos ya sea en el espacio original de los ejemplos de entrada, si éstos son separables, o en un espacio transformado (espacio de características), si los ejemplos no son separables linealmente en el espacio original. La búsqueda del hiperplano de separación en estos espacios transformados, normalmente de muy alta dimensión, se hará empleando las "funciones *kernel*".

Mientras la mayoría de las técnicas pretenden disminuir los errores ejecutados por el modelo generado a partir de los ejemplos de entrenamiento (error empírico), el sesgo inductivo relacionado con las SVM se establece en la minimización del riesgo estructural.

El concepto es hallar un hiperplano de separación que equidista de los ejemplos más próximos de cada clase para obtener un margen máximo a cada lado del hiperplano considerando los ejemplos de entrenamiento (vectores de soporte) de cada clase que se sitúan justo en el límite de dichos márgenes (150).

Desde el punto de vista práctico, el hiperplano separador de margen máximo ha demostrado tener una buena capacidad de generalización, evitando el sobreajuste a los ejemplos de entrenamiento. Se puede ver el concepto de hiperplano óptimo de separación en la Figura 7.



*Figura 7.* Hiperplano óptimo de separación

Desde el punto de vista algorítmico, el problema de optimización del margen geométrico evidencia un problema de optimización cuadrático con restricciones lineales que puede ser solucionado a través de técnicas estándar de programación cuadrática. La convexidad requerida para su resolución reconoce una solución única, en contraste con la no unicidad de la solución hecha por una red neuronal artificial entrenada con un mismo conjunto de ejemplos.

Es importante destacar que el diseño del modelo de la SVM no necesita de toda la totalidad de puntos disponibles para hallar una solución al problema de maximización de la separación entre clases, por lo que el tiempo de entrenamiento no supera 1 segundo en el peor de los casos, lo cual presenta una ventaja frente de otros métodos de clasificación

que utilizan un alto porcentaje del conjunto de entrada (o entrenamiento) para poder hallar una separación entre clases.

Para entender cómo funcionan las SVM, el caso más sencillo que se puede analizar es aquel donde el conjunto de datos de entrenamiento o *inputs* del modelo son linealmente separables. Un ejemplo de este caso se puede representar a partir de un problema con dos variables de entrada y una salida.

### Caso linealmente separable

En un conjunto separable de ejemplos  $S = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)\}$ , cada punto de entrenamiento  $\mathbf{x}_i \in \mathcal{R}^N$  corresponde a una de las dos clases, dándole una etiqueta  $y_i \in \{+1, -1\}$  para  $i = 1, \dots, l$ .

Generalmente, se debe “mapear” el espacio de entrada en un espacio de características de una dimensión mayor para hallar allí el hiperplano óptimo.

Sea  $z = \varphi(x)$  la notación del vector en el espacio de características en un mapeo  $\varphi$  de  $\mathcal{R}^N$  a un espacio de características  $Z$ .

Se pretende hallar el hiperplano

$$\text{(EQ. 8)} \quad w \cdot z + b = 0$$

Definido por el par  $(w, b)$ , tal que se pueda separar el punto  $x_i$  de acuerdo a la función

$$\text{(EQ. 9)} \quad f(\mathbf{x}_i) = \text{sign}(w \cdot z_i + b) = \begin{cases} +1 & y_i = +1 \\ -1 & y_i = -1 \end{cases}$$

Donde  $w \in Z$  y  $b \in \mathcal{R}$ .

Más preciso el conjunto  $S$  es linealmente separable si existe  $(w, b)$  tal que las inecuaciones

$$\text{(EQ. 10)} \quad \begin{cases} (w \cdot z_i + b) \geq +1, & y_i = +1 \\ (w \cdot z_i + b) \leq -1, & y_i = -1 \end{cases} \quad i = 1, \dots, l$$

Sean admitidas para la totalidad de los elementos del conjunto  $S$ . Para el problema linealmente separable de  $S$ , se puede hallar un único hiperplano óptimo, para el cual, el margen es maximizado entre las proyecciones de los puntos de entrenamiento de dos clases distintas.

### Caso no linealmente separable

Para manejar datos que no son linealmente separables, el análisis previo puede ser generalizado introduciendo variables no negativas  $\xi_i \geq 0$  de tal modo que (EQ. 10) es transformada en

$$\text{(EQ. 11)} \quad y_i (w \cdot z_i + b) \geq +1 - \xi_i, \quad i = 1, \dots, l$$

Los  $\xi_i \neq 0$  en (EQ.11) son aquellos para los cuales el punto  $x_i$  no satisface (EQ. 10). El término  $\sum_{i=1}^l \xi_i$  puede ser considerado algún tipo de medida del error en la clasificación. El problema del hiperplano óptimo es rediseñado como la solución al problema.

$$\text{(EQ. 12)} \quad \min \left\{ \frac{1}{2} w \cdot w + C \sum_{i=1}^l \xi_i \right\}$$

$$\text{s.a. } y_i (w \cdot z_i + b) \geq +1 - \xi_i$$

$$\xi_i \geq 0, \quad i = 1, \dots, l$$

$C$  es una constante que puede ser considerada como una medida de regularización, ya que es el único parámetro libre de ser ajustado en la formulación de la SVM, permitiendo hacer un balance entre la maximización del margen y la violación a la clasificación.

Buscar el hiperplano óptimo en (EQ.12) es un problema de programación cuadrática, que puede ser solventado construyendo un Lagrangiano y transformándolo en el dual

$$\text{(EQ. 13)} \quad \text{Max } W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j z_i \cdot z_j$$

$$\text{s.a. } \sum_{i=1}^l y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, \dots, l$$

Donde  $\alpha = (\alpha_1, \dots, \alpha_l)$  es un vector de multiplicadores de Lagrange positivos asociados con las constantes en (EQ.11).

El teorema de Khun-Tucker es significativo en la teoría de las SVM ya que dictamina que la solución  $\bar{\alpha}_i$  al problema (EQ.13) es:

$$\text{(EQ. 14)} \quad \bar{\alpha}_i (y_i (\bar{w} \cdot z_i + \bar{b}) - 1 + \bar{\xi}_i) = 0, \quad i = 1, \dots, l$$

$$\text{(EQ. 15)} \quad (C - \bar{\alpha}_i) \xi_i = 0, \quad i = 1, \dots, l$$

De aquí se concluye que los únicos valores  $\bar{\alpha}_i \neq 0$  (EQ.15) son aquellos que para las constantes en (4) son satisfechas con el signo de igualdad. El punto  $x_i$  correspondiente con  $\bar{\alpha}_i > 0$  es llamado vector de soporte.

Pero existen dos tipos de vectores de soporte en un caso no separable. En el caso  $0 < \bar{\alpha}_i < C$ , el vector de soporte  $x_i$  satisface las igualdades  $y_i(\bar{w} \cdot z_i + \bar{b}) = 1$  y  $\xi_i = 0$ . En el caso  $\bar{\alpha}_i = C$ , el  $\xi_i$  es distinto de cero y el vector de soporte  $x_i$  no satisface (3). Estos vectores de soporte son referidos como errores. El punto  $x_i$  correspondiente con  $\bar{\alpha}_i = 0$ .

Para construir el hiperplano óptimo  $\bar{w} \cdot z_i + \bar{b}$ , se emplea

$$\text{(EQ. 16)} \quad \bar{w} = \sum_{i=1}^l \bar{\alpha}_i y_i z_i$$

#### Truco de Kernel para el caso no linealmente separable

Como no hay ningún conocimiento de  $\varphi$ , el cálculo del problema (EQ.13) y (EQ.16) es imposible; pero en SVM no es necesario conocer  $\varphi$ , sólo se requiere una función  $K(\cdot, \cdot)$  llamada *kernel* que estime el producto punto de los puntos de entrada en el espacio de características  $Z$ , esto es:

$$\text{(EQ. 17)} \quad z_i \cdot z_j = \varphi(x_i) \cdot \varphi(x_j) = K(x_i, x_j)$$

Los kernel codifican y miden la semejanza entre objetos (151–153).

Las funciones que satisfacen el teorema de Mercer (154) pueden ser usadas como productos punto, por lo que pueden ser utilizadas como *kernels*. Se puede emplear el kernel polinomial de grado  $d$

$$\text{(EQ. 18)} \quad K(x_i, x_j) = (1 + x_i \cdot x_j)^d$$

para construir un clasificador SVM. Por lo que el hiperplano no lineal de separación puede ser hallado como la solución de

$$\text{(EQ. 19)} \quad \text{Max } W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$\text{s.a.} \quad \sum_{i=1}^l y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, \dots, l$$

y la función de decisión es

$$\text{(EQ. 20)} \quad f(x) = \text{sign}(w \cdot z + b) = \text{sign}\left(\sum_{i=1}^l \alpha_i y_i K(x_i, x_j) + b\right)$$

#### 3.2.2.4.2 *Random Forest*

---

Random Forest (RF) (155) es un algoritmo para clasificación y regresión de amplio uso que tiene un rendimiento especialmente bueno para datos de alta dimensionalidad. Puede considerarse la integración de las siguientes técnicas: *Decision Trees*, *Bagging* y *Random Subspace*.

Se trata de crear aleatoriamente (seleccionando con reemplazo) un conjunto de entrenamiento del mismo tamaño que el conjunto original. Al elegirse con reemplazo no todos los datos del conjunto general están en el conjunto de entrenamiento (siendo la probabilidad de que esté de un 66%).

Los datos que no forman parte del conjunto de entrenamiento forman el conjunto de validación o *out of bag data* (en adelante, OOB data).

En cada punto de división del árbol o nodo, la averiguación de la mejor variable para dividir los datos se realiza sobre un subconjunto  $m$  de variables elegido aleatoriamente, buscando la mejor división de los datos de entrenamiento contabilizando sólo a las  $m$  variables seleccionadas e implementando para ello una función objetivo que habitualmente es la entropía o el “índice de Gini”.

Los pasos previos son reproducidos en varias ocasiones, dando lugar a un conjunto de árboles de decisión entrenados sobre distintos conjuntos de datos y de atributos, dejando crecer el árbol en profundidad; es decir, sin aplicar sobre el mismo técnicas de poda de posibles soluciones.

Posteriormente, la estimación de cada nueva entrada es ejecutada con el conjunto de árboles. La categoría final de la clase (clasificación) es realizada por el voto mayoritario del conjunto de árboles, y en el caso de regresión por el valor promedio de los resultados.

Para medir el error de RF se suele utilizar la técnica denominada “OOB error “. Para cada árbol, se utiliza el conjunto de objetos no seleccionados por su muestra *bootstrap* de entrenamiento para ser clasificados en dicho árbol. Promediando sobre el conjunto de árboles de RF se puede estimar el error del algoritmo.

Además, RF puede ser paralelizado eficazmente puesto que cada árbol puede construirse de manera independiente a los otros, siendo deseable para entornos paralelos.

### Importancia de las variables

El mecanismo de construcción de RF permite establecer un baremo de la importancia de cada variable en la predicción final. Para ello, se calcula el error de la muestra OOB. A continuación, para cada variable de la muestra, se permutan un par de elementos y se vuelve a calcular el error de la muestra OOB permutada. El resultado debería ser peor que para la muestra original.

Este procedimiento se realiza para todos los valores de todas las variables y se calcula el promedio. Así, las variables menos importantes deberían alterar menos la diferencia entre el error de la muestra original y la permutada.

#### 3.2.2.4.3 K vecinos más próximos (kNN)

El kNN es un algoritmo de aprendizaje supervisado, el cual, a partir de un conjunto de datos inicial se clasificarán correctamente todas las instancias nuevas. El dataset típico está compuesto por varios atributos descriptivos y un solo atributo objetivo (también llamado clase).

Comparado con otros algoritmos de aprendizaje supervisado, kNN no genera un modelo fruto del aprendizaje con datos de entrenamiento. Los pronósticos se basan en los ejemplos más similares al que hay que predecir, por lo que el coste de aprendizaje reside en el cálculo de la predicción. A este tipo de algoritmos se les llama *lazy learning methods*.

Se debe detallar una métrica para medir la proximidad, siendo habitual la distancia Euclídea,  $\delta$ . El conjunto de referencia es el conjunto de prototipos sobre el que se buscará el/los vecino/s más cercano/s (156).

#### Regla 1-NN

La regla más sencilla es la de clasificación del vecino más cercano o simplemente 1- NN, basada en la hipótesis de que la clase del patrón a etiquetar,  $X$ , es la del prototipo más cercano en  $R$ , (conjunto de referencia) que será  $X_{NN}$ . Esta regla puede formularse de la siguiente manera:

$$(EQ. 21) \quad d(X) = w_c \text{si} \delta(X, X_{NN}) = \min_{i=1-N} \delta(X, X_i)$$

En la Figura 8, se muestra cómo se clasificaría el patrón X con la regla 1NN para un problema de clasificación de dos clases. Hay cuatro prototipos de clase 1 y cinco prototipos de clase 2. El prototipo más próximo es de clase 1, por lo que esta será la clase asociada a X.

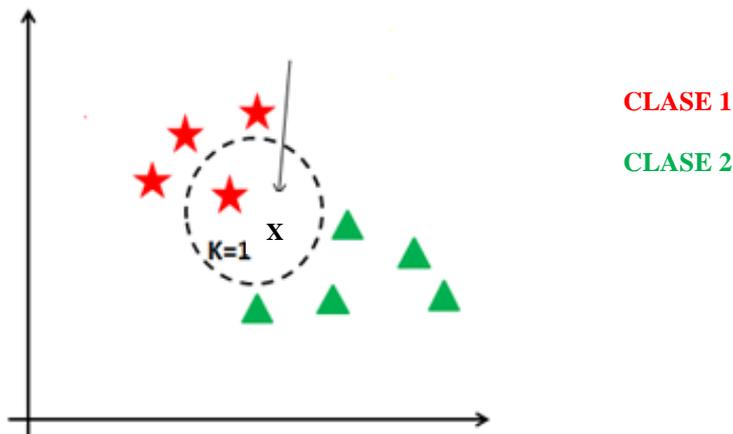


Figura 8. Clasificación del patrón X con 1NN

### Regla k-NN

La regla más frecuente es la regla de clasificación de los k vecinos más cercanos o k-NN. Se basa en que los prototipos más cercanos tienen una probabilidad a posteriori semejante. Si  $K_i(X)$  es el número de muestras de la clase presentes en los k vecinos más próximos a X. Esta regla puede formularse como:

$$(EQ. 22) \quad d(X) = w_c \text{si} K_c(X) = \max_{i=1 \rightarrow J} K_i(K)$$

A continuación, se describen las fases para diseñar este algoritmo:

Durante la etapa de aprendizaje se unen en un vector el conjunto de entrenamiento, junto a la clase asociada a cada muestra de este conjunto.

Primero, se calcula la distancia euclídea de cada muestra de entrenamiento y las que están almacenadas en el vector del punto anterior, quedando las K muestras más próximas y clasificando la nueva muestra de entrenamiento en la clase más frecuente a la que pertenecen los K vecinos obtenidos anteriormente.

Posteriormente, se lleva a cabo el mismo proceso con los datos de validación, estimando el porcentaje de clasificación sobre los ejemplos de este conjunto (desconocidos en la tarea de aprendizaje) para descubrir su poder de generalización.

### Elección de K

En un espacio bidimensional con varios prototipos de una misma clase representados en él, con un patrón cualquiera  $X$ , considerando los  $k$  prototipos más próximos a  $X$ , éstos estarán ubicados en un círculo centrado en  $X$ .

En espacios multidimensionales, el círculo se transforma en una hiperesfera, Si  $K_i(X)$   $i = 1, 2, \dots, J$  son los prototipos de clase  $w_i$  que se localizan en una vecindad de  $X$ , el valor de

$$(EQ. 23) \quad \hat{p}(X|w_i)$$

puede calcularse como la proporción de los  $N_i$  prototipos de la clase que están en esa vecindad:

$$(EQ. 24) \quad \hat{p} = \frac{K_i(X)}{N_i}$$

Teniendo en cuenta que, en una región muy poblada, la vecindad tendrá un radio menor, y poco poblada, un radio superior, habrá que cambiar la ecuación anterior de forma que se pondere inversamente con el área de la región considerada. Por lo que se obtiene:

$$(EQ. 25) \quad \hat{p} = \frac{K_i(X)}{N_i V(X)}$$

La cuestión que se plantea es si hay un valor óptimo para  $k$  o si se pueden dar algunas reglas para fijarlo. La elección de  $k$  está determinada por la densidad de los puntos y debería realizarse en función a  $N_i$ . Se puede probar que el estimador (el anterior) es insesgado y sólido si se cotejan las siguientes condiciones sobre  $k$ :

$$(EQ. 26) \quad \lim_{N_i \rightarrow \infty} k(N_i) = \infty$$

$$\lim_{N_i \rightarrow \infty} \frac{k(N_i)}{N_i} = \infty$$

Así, una opción adecuada de  $k(N_i)$  es:  $N_i = cte \sqrt{N_i}$

Por todo ello, se concluye que son dos los factores que establecen el coste computacional de las reglas k-NN: la dimensionalidad de los datos,  $d$ ; y el tamaño del conjunto de referencia,  $N$ .

Para ello, prevalecen dos clases de estrategias que permiten el empleo de la regla del vecino más cercano para múltiples conjuntos de referencia con un coste computacional menor que el asociado a la fuerza bruta (157):

- Reducir el conjunto de referencia: Eligiendo un subconjunto del conjunto de referencia con las mismas propiedades que el conjunto original para, una vez reducido, aplicar la regla del vecino más cercano en su formulación original (fuerza bruta) sobre el nuevo conjunto.
- Optimizar la eficiencia computacional de la búsqueda del vecino más cercano: trabajando con el conjunto completo de referencia, su objetivo es disminuir el número de cálculos, excluyendo los inútiles o innecesarios. Entre estos métodos están los basados en ordenaciones y los jerárquicos.

Los métodos basados en ordenaciones clasifican los patrones del conjunto de referencia en función de alguna coordenada, y la búsqueda se hace de acuerdo con alguna propiedad de la métrica asociada al espacio. Los métodos jerárquicos distribuyen los patrones en una estructura de tipo árbol, para posteriormente aplicar una técnica de poda y ramificación (*branch and bound*) para explorar el árbol.

### 3.2.2.4.4 Modelos lineales generalizados

#### Regresión Ridge

Este método también conocido como regresión contraída o *Tikhonov regularization* fue propuesto por Hoerl y Kennard (79) para sortear las adversidades que conlleva de la colinealidad en un modelo lineal estimado por mínimos cuadrados, en el contexto  $p < n$ .

Regresión Ridge es muy parecido a los mínimos cuadrados, excepto que los coeficientes se estiman minimizando una cantidad distinta.

Los coeficientes estimados por Ridge,  $\hat{\beta}^{ridge}$ , son los valores que minimizan

$$(EQ. 27) \quad \hat{\beta}^{ridge} = \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j X_{ij})^2 + \lambda \sum_{j=1}^p \beta_j^2 = RSS + \lambda \sum_{j=1}^p \beta_j^2$$

Donde  $\lambda \geq 0$  es el parámetro de contracción que se establecerá por separado.

El método Ridge aplica una penalización al tamaño de los coeficientes de la relación lineal entre las características predictivas y la variable objetivo: cuanto mayor sea  $\lambda$ , mayor penalización y, por tanto, mayor contracción de los coeficientes.

Una forma similar de formular el problema Ridge es:

$$(EQ. 28) \quad \hat{\beta}^{ridge} = \arg \min_{\beta} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p X_{ij} \beta_j)^2$$

$$\text{s.a. } \sum_{j=1}^p \beta_j^2 \leq t,$$

Donde  $t$  es el parámetro de penalización por complejidad.

$\hat{\beta}^{MCO} = (X'X)^{-1}X'Y$  es la estimación por mínimos cuadrados de  $\beta$ . La escasa estabilidad de  $\hat{\beta}^{MCO}$  puede ser mejorada añadiendo una constante,  $\lambda \geq 0$ , a cada término de la diagonal de  $(X'X)$  antes de volcar la matriz:

$$(EQ. 29) \quad \hat{\beta}^{ridge}(\lambda) = (X'X + \lambda I_p)^{-1}X'Y,$$

Donde  $I$  es la matriz  $p \times p$ .

Dicha penalización se aplica a los coeficientes  $\beta_1, \dots, \beta_p$  pero no a  $\beta_0$ . Se ajusta el modelo sin término independiente, estimando éste mediante  $\bar{y} = \sum_{i=1}^n \frac{y_i}{n}$ . Es exceptuado de la penalización para que el resultado no dependa del origen en la variable Y.

A diferencia de mínimos cuadrados, regresión Ridge origina un conjunto distinto para cada valor de  $\lambda$  y no únicamente vector de coeficientes estimados.

Si  $\lambda = 0$  es un caso de mínimos cuadrados ordinarios. En otro caso,  $\lambda \rightarrow \infty, \hat{\beta} \rightarrow 0$  es un estimador sesgado de  $\beta$ . Es decir, se incorpora sesgo, pero se reduce la varianza.

Para que la penalización no cambie frente a modificaciones de escalas de las variables, estas son previamente estandarizadas (media 0 y varianza 1). También Y está centrada, por lo que la matriz X tendrá  $p$  columnas y no  $p + 1$ .

Una vez estimados los coeficientes, a través de la validación cruzada, se busca el valor de  $\lambda, 0 < \lambda < \infty$ , para minimizar una estimación del error de predicción esperado.

Una de las desventajas de esta técnica es que contrae todos los coeficientes hacia cero, pero sin obtener la nulidad de ninguno, por lo que no se da una selección de variables, siendo un inconveniente en aquellos estudios que tienen una gran cantidad de variables explicativas o predictoras.

Para evitar este inconveniente se plantea la regresión Lasso, explicada a continuación.

### Regresión Lasso

El método Lasso (*least absolute shrinkage and selection operator*), propuesto por Tibshirani (1996), es una técnica de regresión lineal regularizada, como Ridge, pero que, a través de una selección de variables contrae los coeficientes, estabilizando las estimaciones y predicciones.

A partir de un determinado valor del parámetro de complejidad, el estimador de Lasso realiza estimaciones nulas para ciertos coeficientes y no nulas para otros, realizando una selección de variables en forma continua, debido a la norma L1.

Lasso disminuye la variabilidad de las estimaciones al acortar el número de coeficientes, a la par que produce modelos interpretables debido a la reducción de determinados coeficientes a cero.

Actualmente, los problemas donde  $p \gg n$  y el desarrollo paralelo de algoritmos eficientes (80), han aumentado el empleo de técnicas Lasso.

Lasso solventa el problema de mínimos cuadrados con limitación sobre la norma L1 del vector de coeficientes:

$$(EQ. 30) \quad \hat{\beta}^{\text{lasso}} = \min_{\beta} \left\{ \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j X_{ij})^2 \right\}$$

$$\text{s.a. } \sum_{j=1}^p |\beta_j| \leq s$$

o de manera similar, minimizando

$$(EQ. 31) \quad \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j X_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j| = \text{RSS} + \lambda \sum_{j=1}^p |\beta_j|$$

siendo  $s$  y  $\lambda \geq 0$  los correspondientes parámetros de penalización por complejidad.

Nuevamente, se reparametriza la constante  $\beta_0$  a través de la estandarización de los predictores, y, por tanto, la solución  $\hat{\beta}_0 = \bar{y}$ .

Los modelos generalizados Lasso son mucho más sencillos de interpretar que los realizados con Ridge. Finalmente, queda por encontrar el mejor valor del parámetro  $\lambda$  a través de la validación cruzada.

Sin embargo, esta técnica también tiene ciertas desventajas:

- En el problema  $p > n$ , Lasso elige como máximo  $n$  variables antes de saturarse, debido al problema de optimización convexa. Por otro lado, Lasso no es concreto a menos que el límite de la norma L1 de los coeficientes sea más pequeño que un determinado valor.
- Si existe un conjunto de variables con altas correlaciones por parejas, Lasso elige sólo una variable del grupo, sin importarle cuál.
- En el caso  $n > p$ , con alta correlación entre variables, generalmente, la predicción a través de Ridge obtiene mejores resultados que a través de Lasso.

Por todo ello, haciendo una comparativa, se puede concluir que:

- La regresión Lasso al realizar una selección de variables ofrece una importante ventaja sobre Ridge, obteniendo modelos más sencillos e interpretables con un único subconjunto de predictores. Sin embargo, no hay una técnica que siempre garantice ser la mejor.
- Puede ser que Lasso dé mejores resultados en un caso con un número pequeño de predictores con coeficientes sustanciales, y con los restantes predictores con coeficientes escasos o iguales a cero.

- La regresión Ridge es mejor cuando la respuesta es una función de múltiples factores predictivos, todos con coeficientes de tamaño similar. Se debe tener en cuenta que, para los conjuntos de datos reales, se desconoce previamente el número de predictores relacionados con la respuesta.
- Para estipular el mejor enfoque en un conjunto de datos específico, se emplea la validación cruzada.
- Al igual que la regresión Ridge, cuando las estimaciones de mínimos cuadrados tiene excesivamente alta varianza, la solución Lasso puede producir una reducción de la varianza a expensas de un pequeño aumento de sesgo, y por tanto, puede generar predicciones más exactas al estar el algoritmo sobreentrenado.

### Elastic Net

Este método propuesto por Zou y Hastie (78), es una técnica de regularización y selección de variables que mantiene las ventajas de Lasso y, a mayores, supera algunas de sus limitaciones, ya que permite seleccionar grupos de variables correlacionadas.

Este método es especialmente útil cuando el número de predictores ( $P$ ) es mucho más grande que el número de observaciones ( $n$ ).

Inicialmente, “Naive Elastic Net” (red elástica simple), es un método de mínimos cuadrados con una penalización propia de Elastic Net.

### **Naive Elastic Net (78)**

Para cualquiera  $\lambda_1$  y  $\lambda_2$  constantes fijas no negativas, se formula el criterio “Elastic Net” simple por:

$$\text{(EQ. 32)} \quad L(\lambda_1, \lambda_2, \beta) = |y - X\beta|^2 + \lambda_2 |\beta|^2 + \lambda_1 |\beta|_1$$

Donde  $|\beta|^2 = \sum_{j=1}^p \beta_j^2$ ,  $|\beta|_1 = \sum_{j=1}^p |\beta_j|$  y  $|y - X\beta|^2 = RSS$

El estimador obtenido, es decir, el vector que minimiza la ecuación anterior

$$\text{(EQ. 33)} \quad \hat{\beta}^{\text{ene}} = \arg \min_{\beta} L(\lambda_1, \lambda_2, \beta)$$

se denomina estimador “Elastic Net” simple.

Suponiendo  $\alpha = \frac{\lambda_2}{\lambda_1 + \lambda_2}$ , se obtiene el consiguiente problema de optimización:

$$\text{(EQ. 34)} \quad \hat{\beta}^{\text{ene}} = \arg \min_{\beta} |y - X\beta|^2$$

s.a.:  $\alpha |\beta|_1 + (1 - \alpha)|\beta|^2 \leq t$  para algún  $t$ .

Se define a la función  $\alpha |\beta|_1 + (1 - \alpha)|\beta|^2$  como la penalización Elastic Net, que es una combinación convexa de las penalizaciones Lasso y Ridge.

### Elección del parámetro $\lambda$

Como se ha visto, todas estas técnicas de mínimos cuadrados penalizados dependen de un parámetro de penalización  $\lambda$ , que controla la importancia dada a la penalización en el proceso de optimización. Cuanto mayor es  $\lambda$  mayor es la penalización en los coeficientes de regresión y más son contraídos éstos hacia cero.

La elección de este parámetro involucra un balance entre los componentes de sesgo y varianza del error cuadrático medio al estimar  $\beta$ .

Una propuesta inicial, y que continúa siendo sugerida por algunos autores, es la utilización de una traza Ridge para determinar  $\lambda$ . Consiste en representar simultáneamente los coeficientes de regresión estimados en función de  $\lambda$ , y elegir el valor más pequeño del parámetro para el cual se estabilizan dichos coeficientes.

Un método más automático, pero intensivo computacionalmente, consiste en estimar  $\lambda$  mediante validación cruzada (en general se recomiendan utilizar ambos métodos y comparar resultados).

El valor del parámetro seleccionado será el que ofrezca el mínimo error.



## 4. RESULTADOS



## 4.1. CARACTERÍSTICAS GENERALES DE LA MUESTRA ESTUDIADA

### 4.1.1. VARIABLES SOCIODEMOGRÁFICAS

Se incluyeron en el estudio n=503 personas de 40 y más años. Las características generales de la muestra se observan en la Tabla 11. La media de la edad es de 63,1±12,1 años, con una mediana de 65. Existe un predominio del sexo femenino (59%) como corresponde en la distribución poblacional del municipio de Cambre (A Coruña) dónde se recogieron los datos del trabajo de campo. La mayoría de la población refiere estudios de primer grado (54,8%). Un 78,1% están casados o en pareja y viviendo solos el 7,5%.

Tabla 11. *Distribución de la muestra estudiada según variables sociodemográficas.*

	<b>n</b>	<b>Media ± DT</b>	<b>Mediana</b>
<b>Edad (años)</b>	503	63,1±12,1	65,0
	<b>n</b>	<b>%</b>	<b>95%IC</b>
<b>Sexo</b>			
Hombre	206	41,0	36,6-45,3
Mujer	297	59,0	54,6-63,4
<b>Nivel de estudios</b>			
Sin estudios	91	18,1	14,6-21,5
1º grado incompleto	223	44,4	39,9-48,7
1º grado (EGB)	52	10,4	7,6-13,1
2º grado	87	17,3	14,1-20,9
Universitario	49	9,8	7,0-12,4
<b>Estado civil</b>			
Soltero	23	4,7	2,7-6,6
Casado/pareja	385	78,1	74,3-81,8
Viudo	63	12,8	9,7-15,8
Divorciado	12	2,4	0,9-3,9
Separado	9	1,8	0,5-3,1
<b>¿Vive solo?</b>			
No	455	92,5	90,0-94,9
Si	37	7,5	5,1-9,9

Tras estratificar la muestra en dos grupos de edad (40-64 años vs 65 y más años) se observa que existen diferencias significativas entre ambos grupos en relación nivel de estudios y el hecho de vivir sólo. El nivel de estudios es significativamente más elevado en el grupo más joven. En el grupo de mayor edad, se objetiva que viven solos con mayor frecuencia (9,7% vs 4,9%) (Tabla 12).

Tabla 12. *Características sociodemográficas de la muestra estudiada, según grupos de edad.*

	40-64 años		65 y más		p
	n (%)	Media±DT (Mediana)	n (%)	Media±DT (Mediana)	
<b>Edad</b>	233(46,3)	51,7 ±5,8(50,0)	270(53,7)	72,9±6,1(72,0)	<b>&lt;0,001</b>
<b>Sexo</b>					0,774
Hombre	97(41,6)		109(40,4)		
Mujer	136(58,4)		161(59,6)		
<b>Nivel de estudios</b>					<b>&lt;0,001</b>
Sin estudios	1(0,4)		90(33,5)		
1º grado incompleto	76(32,6)		147(54,6)		
1º grado (EGB)	48(20,6)		4(1,5)		
2º grado	73(31,3)		14(5,2)		
Universitario	35(15,0)		14(5,2)		
<b>Estado civil</b>					<b>&lt;0,001</b>
Soltero	15(6,6)		8(3,0)		
Casado/pareja	187(83,1)		198(73,8)		
Viudo	7(3,1)		56(21,0)		
Divorciado	9(4,0)		3(1,1)		
Separado	7(3,1)		2(0,7)		
<b>¿Vive sólo?</b>					<b>0,042</b>
No	214(95,1)		241(90,3)		
Si	11(4,9)		26(9,7)		

#### 4.1.2. COMORBILIDAD Y HÁBITO TABÁQUICO

Los datos de comorbilidad se muestran en la Tabla 13. En dicha tabla se observa que la mediana del Score de Charlson es de 2 y que la patología más prevalente de las identificadas de este Score es la diabetes mellitus (12,8%).

Tabla 13. *Comorbilidad, tensión arterial y hábito tabáquico.*

	n	%	95%IC	Media±DT	Mediana
Score de Charlson	469			0,5±0,9	0,0
Score de Charlson (ajustado por edad)	469			2,3±1,8	2,0
<b>Comorbilidades (Score de Charlson)</b>					
SIDA (no únicamente VIH positivo)	0	0	0-0		
Enfermedad Cerebrovascular	9	1,9	0,6-3,3		
EPOC (Enfermedad Pulmonar Obstructiva Crónica)	27	5,7	3,5-7,9		
Insuficiencia Cardíaca Congestiva	7	1,5	0,3-2,7		
Enfermedad de tejido conectivo	8	1,7	0,4-2,9		
Demencia	4	0,8	0,2-2,2		
Hemiplejía	0	0	0-0		
Leucemia (aguda o crónica)	0	0	0-0		
Linfoma	4	0,8	0,2-2,2		
Infarto de miocardio	22	4,7	2,7-6,7		
Enfermedad vascular periférica	11	2,3	0,9-3,8		
Úlcus péptico	27	5,7	3,5-7,9		
Diabetes Mellitus con afectación de órganos	7	1,5	0,3-2,7		
Diabetes Mellitus sin afectación de órganos	53	11,3	8,3-14,3		
Enfermedad hepática leve	12	2,6	1,0-4,1		
Enfermedad hepática moderada	1	0,2	0,005-1,2		
Enfermedad hepática severa	3	0,6	0,1-1,8		
Enfermedad renal leve	5	1,1	0,3-2,5		
Enfermedad renal moderada	9	1,9	0,6-3,3		
Tumor sólido maligno (sin metástasis)	23	4,9	2,8-6,9		
Tumor sólido maligno (con metástasis)	1	0,2	0,005-1,2		

CONTINUACIÓN TABLA 13

	n	%	95%IC	Media±DT	Mediana
<b>TAS (mmHg)</b>	492			128,2±14,9	125,0
<b>TAS ≥140mmHg</b>	143	29,0	24,9-33,2		
<b>TAD (mmHg)</b>	492			74,4±10,1	70,0
<b>TAD ≥90mmHg</b>	43	8,7	6,1-11,3		
<b>HTA (TAS≥140mmHg y TAD≥90mmHg)</b>	152	30,2	26,7-35,1		
<b>Antecedentes de HTA</b>					
Si	185	37,9	33,5-42,3		
No	303	62,1	57,7-66,5		
<b>Hábito tabáquico</b>					
No fumador	305	60,8	56,4-65,1		
Exfumador	127	25,3	21,4-29,2		
Fumador	70	13,9	10,8-17,1		
<b>Test de Fagerström</b>	64			7,1±2,5	7,0
Dependencia baja	6	9,4	1,4-17,3		
Dependencia moderada	31	48,4	35,4-61,5		
Dependencia alta	27	42,2	29,3-55,1		

Tras la medición de la tensión arterial, presentan HTA (TAS≥140mmHg y TAD≥90mmHg) el 30,2% de los casos en el momento de la medición. En la historia clínica figura con antecedentes de HTA el 37,9%. La prevalencia de tabaquismo es de un 13,9%, presentando dependencia moderada el 48,4% (Tabla 14).

Por edad, en relación a la comorbilidad existen diferencias significativas entre los dos grupos de edad establecidos en: el índice de Charlson, la prevalencia de HTA y la prevalencia de tabaquismo. El grupo de 65 años o más tiene un índice de Charlson sin ajustar por edad discretamente mayor ( $0,7 \pm 1,2$  vs  $0,2 \pm 0,6$ ), mayor prevalencia de HTA objetivada en el momento de la medición (36,9% vs 22,4%) y menor prevalencia de tabaquismo (4,5% vs 24,9%) (Tabla 14).

Tabla 14. Comorbilidad y hábito tabáquico según grupos de edad

	40-64 años		65 y más		p
	n (%)	Media±DT (Mediana)	n (%)	Media±DT (Mediana)	
Score de Charlson	218	0,2±0,6 (0)	251	0,7 ±1,2(0)	<0,001
Score de Charlson (ajustado por edad)	218	0,9 ± 0,9 (1,0)	251	3,5±1,5(3,0)	<0,001
<b>Comorbilidades (Score de Charlson)</b>					
SIDA (no sólo VIH positivo)	0		0		
Enfermedad Cerebrovascular	0		9(3,5)		
EPOC (Enfermedad Pulmonar Obstructiva Crónica)	7 (3,2)		20(7,9)		
Insuficiencia Cardiaca Congestiva	1(0,5)		6(2,4)		
Enfermedad de tejido conectivo	3(1,4)		5(2,0)		
Demencia	0		4(1,6)		
Hemiplejia	0		0		
Leucemia (aguda o crónica)	0		0		
Linfoma	1(0,5)		3(1,2)		
Infarto de miocardio	6(2,7)		16(6,3)		
Enfermedad vascular periférica	2(0,9)		9(3,6)		
Úlcus péptico	6(2,7)		21(8,3)		
Diabetes Mellitus					
Con afectación de órganos	1(0,5)		6(2,4)		
Sin afectación de órganos	15(6,8)		38(15,1)		
Enfermedad hepática					
Leve	8(3,7)		4(1,6)		
Moderada	0		1(0,4)		
Severa	0		3(1,2)		
Enfermedad renal					
Leve	1(0,5)		4(1,6)		
Moderada	0		9(3,6)		
Tumor sólido maligno					
Sin metástasis	4(1,8)		19(7,6)		
Con metástasis	0		1 (0,4)		

<b>CONTINUACIÓN TABLA 14</b>					
	<b>n(%)</b>	<b>40-64 años Media±DT (Mediana)</b>	<b>n(%)</b>	<b>65 y más Media±DT (Mediana)</b>	<b>p</b>
<b>TAS (mmHg)</b>	232	124,6±15,3 (122,5)	260	131,4±13,8 (260)	<b>&lt;0,001</b>
TAS ≥140mmHg	45(19,4)		98(37,7)		
<b>TAD (mmHg)</b>	232	74,8 ± 10,1 (70)	260	74,2±10,2 (75,0)	0,067
TAD ≥90mmHg	26(11,2)		17(6,5)		
<b>HTA (TAS≥140mmHg y TAD≥90mmHg)</b>	52(22,4)		100(36,9)		<b>&lt;0,001</b>
<b>Antecedentes de HTA</b>					<b>&lt;0,001</b>
Si	41(18,5)		144(54,1)		
No	181(81,5)		122(45,9)		
<b>Hábito tabáquico</b>					<b>&lt;0,001</b>
No fumador	109(46,8)		196(72,9)		
Exfumador	66(28,3)		61(22,7)		
Fumador	58(24,9)		12(4,5)		
<b>Test de Fagerström</b>	56	7,0±2,6 (7,0)	8	7,3±2,2 (6,0)	0,528
Dependencia baja	6(10,7)		0		
Dependencia moderada	26(46,4)		5(62,5)		
Dependencia alta	24(42,9)		3(37,5)		

#### 4.1.3. CARACTERÍSTICAS ANTROPOMÉTRICAS

La prevalencia de sobrepeso es del 45,7% y de obesidad del 37,7% (Tabla 15). La mediana del índice cintura-cadera es de 0,9 y las diferentes mediciones de los perímetros, pliegues, área muscular, área grasa del brazo y bioimpedancias se muestran en la Tabla 15.

Tabla 15 .Características antropométricas de la muestra estudiada.

	n (%)	95%IC	Media±DT	Mediana	Rango
<b>Peso (Kg)</b>	502		74,6±13,9	73,6	
<b>Talla (cm)</b>	503		160,2±9,5	160,0	
<b>Índice de masa corporal(kg/m<sup>2</sup>)</b>			29,02±4,7	28,4	
Normopeso (IMC 18,5-24,9Kg/m <sup>2</sup> )	83(16,6)	13,2-19,9			
Sobrepeso (IMC 25-29, 9Kg/m <sup>2</sup> )	229(45,7)	41,2-50,2			
Obesidad (IMC ≥ 30Kg/m <sup>2</sup> )	189(37,7)	33,4-42,1			
<b>Perímetro de cintura(cm)</b>	502		93,5±12,2	93,7	64,0-150,0
<b>Perímetro de cadera(cm)</b>	502		102,4±9,7	101,0	80-174
<b>Índice cintura – cadera</b>	502		0,91±0,09	0,9	0,69-1,22
<b>Pliegues cutáneos(mm)</b>					
Tricipital	502		15,7±4,9	16,0	3,0-41,0
Bicipital	502		14,3±5,6	14,0	3,0-55,0
Subescapular	502		19,6±5,4	19,0	5,0-41,0
Suprailíaco	502		18,8±6,5	19,0	3,0-44,0
<b>Perímetro de pantorrilla(cm)</b>	499		34,7±3,3	34,5	27,0-55,0
<b>Perímetro del brazo(cm)</b>	502		30,0 ±3,4	30,0	21,0-49,0
<b>Perímetro muscular del brazo(cm)</b>	502		25,1±3,1	24,8	17,8-36,1
<b>Área muscular del brazo</b>	502		8,2±2,1	7,9	4,1-16,8
<b>Área grasa del brazo</b>	502		3,5±1,3	3,5	0,6-14,1
<b>Índice adiposo muscular</b>	502		0,4±0,2	0,4	0,07-1,1
<b>Bioimpedancia</b>					
Masa adiposa	499		33,7 (7,0)	33,9	3,6-49,0
Masa líquida	499		48,3(5,0)	48,2	37,2-64,1
Masa muscular	499		32,0 (5,0)	31,2	23,2-42,9
Metabolismo basal calórico	499		1423,8(228,4)	1369,5	965,0-2214,0
Masa ósea	499		6,9(2,3)	6,6	2,5-13,1
Metabolismo en actividad	499		2063,2(413,4)	1965,5	1196,0-3993,0

Existen diferencias estadísticamente significativas entre los grupos de edad (40-64 años vs 65 y más años) (Tabla 16) en los siguientes índices, perímetros, áreas y mediciones de bioimpedancia: el índice de masa corporal, perímetro de la cintura, perímetro de la cadera, perímetro del brazo, perímetro de pantorrilla, perímetro muscular del brazo, área muscular del brazo, área grasa del brazo y en las mediciones de la bioimpedancia.

Tabla 16. Características antropométricas según grupos de edad

	40-64 años			65 y más			P
	n (%)	Media±DT	Mediana	n (%)	Media±DT	Mediana	
<b>Peso (Kg)</b>	233	75,8 ± 14,7	74,3	269	73,5±13,1	72,9	0,151
<b>Talla (cm)</b>	233	163,7 ± 9,2	163,0	270	157,2±8,7	157,0	<b>&lt;0,001</b>
<b>IMC</b>	230	28,3±4,8	27,8	270	29,7±4,5	29,5	<b>&lt;0,001</b>
Normopeso (IMC 18,5-24,9Kg/m <sup>2</sup> )	46(20,0)			37(13,7)			
Sobrepeso (IMC 25-29, 9Kg/m <sup>2</sup> )	120(52,2)			108(40,0)			
Obesidad (IMC ≥ 30Kg/m <sup>2</sup> )	64(27,8)			125(46,3)			
<b>Perímetro de cintura</b>	232	91,3 ±12,2	91,0	270	95,4±11,8	96,0	<b>&lt;0,001</b>
<b>Perímetro de cadera</b>	232	100,9±9,4	99,0	270	103,7±9,8	103,0	<b>&lt;0,001</b>
<b>Índice cintura - cadera</b>	232	0,91±0,09	0,9	270	0,92±0,08	0,9	0,096
<b>Perímetro del brazo</b>	230	30,6±3,5	30,0	271	29,6±3,3	29,0	<b>0,002</b>
<b>Perímetro de pantorrilla</b>	229	35,4±3,2	35,0	269	34,1±3,2	34,0	<b>&lt;0,001</b>
<b>Pliegues cutáneos</b>							
Tricipital	230	16,2±4,5	16,0	271	15,5±5,2	15,0	0,145
Bicipital	230	14,7±5,7	14,0	271	14,1±5,6	13,0	0,212
Subescapular	230	20,0±5,6	20,0	271	19,3±5,2	19,0	0,166
Supraíliaco	230	19,5±6,2	19,0	271	18,1±6,7	18,0	0,061
<b>Perímetro muscular del brazo</b>	230	25,5±3,4	25,2	271	24,7±2,9	24,5	<b>0,016</b>
<b>Área muscular del brazo</b>	230	8,5±2,2	8,2	271	7,9±1,9	7,7	<b>0,016</b>
<b>Área grasa del brazo</b>	230	3,6±1,3	3,5	271	3,4±1,3	3,4	<b>0,034</b>
<b>Índice adiposo muscular</b>	230	0,4±0,1	0,4	271	0,4±0,2	0,4	0,490
<b>Bioimpedancia</b>							
Masa adiposa	229	31,7±5,9	31,7	269	35,4±7,4	36,4	<b>&lt;0,001</b>
Masa líquida	229	49,8±4,4	49,9	269	47,1±5,2	46,4	<b>&lt;0,001</b>
Masa muscular	229	33,7±4,4	32,0	269	30,5±5,0	27,8	<b>&lt;0,001</b>
Metabolismo basal calórico	229	1517,0±232,9	1459,0	269	1343,9±191,9	1319,0	<b>&lt;0,001</b>
Masa ósea	229	8,3±2,2	7,7	269	5,8±1,9	5,5	<b>&lt;0,001</b>
Metabolismo en actividad	229	2217,1±455,0	2114,0	269	1932,0±322,3	1865,0	<b>&lt;0,001</b>

El índice de masa corporal es significativa y discretamente mayor en el grupo de mayor edad (29,7±4,5 vs 28,3±4,8). La prevalencia de obesidad es significativamente mayor entre los mayores de 65 años (46,3% vs. 27,8%).

Los perímetros de la cintura y la cadera son mayores en los de mayor edad, encontrándose el índice cintura-cadera en el límite de significación estadística. En los más jóvenes se observa que los perímetros de brazo y pantorrilla, el perímetro muscular, el área muscular y el área grasa del brazo son mayores que en el grupo de 65 y más años.

En relación a la bioimpedancia se objetiva que la gente de más edad tiene valores más elevados en la determinación de la bioimpedancia de la masa adiposa. En todas las demás mediciones (masa líquida, masa muscular, metabolismo basal calórico, masa ósea y metabolismo en actividad) los valores son más elevados en la gente más joven (Tabla 16).

#### 4.1.4. AUTOPERCEPCIÓN DE LA IMAGEN CORPORAL

##### 4.1.4.1 Percepción subjetiva del peso

La valoración de la percepción subjetiva de peso se muestra en la Tabla 17, donde se indica que el 64,4% de los participantes consideran que están por encima de su peso.

Tabla 17. *Percepción subjetiva de peso en la muestra estudiada*

Percepción subjetiva de peso	n (%)	95%IC	Total Media±DT	Mediana
<b>Puntuación global</b>			2,6±0,5	3,0
<b>Categorías</b>				
Superior a lo normal	322(64,4)	60,1-68,7		
Normal	162(32,4)	28,2-36,6		
Inferior a lo normal	16(3,2)	1,6-4,8		

Considerando la puntuación global de la percepción subjetiva de peso existe un ligero incremento de dicho valor en la población de 40-64 años en comparación con los de 65 y más. Tras categorizar los valores según considere el individuo si su peso es superior, normal o inferior al normal, existen diferencias estadísticamente significativas entre dichas categorías y los grupos de edad. La creencia de que su peso es superior al normal es mayor en el grupo más joven (69,6% vs 60,0%) (Tabla 18).

Tabla 18. *Percepción subjetiva de peso según grupos de edad*

Percepción subjetiva de peso	40-64 años		65 y más años		p
	Media±DT	Mediana	Media±DT	Mediana	
<b>Puntuación global</b>	2,7±0,5	3,0	2,5±0,6	3,0	<b>0,019</b>
	<b>n</b>	<b>%</b>	<b>n</b>	<b>%</b>	<b>p</b>
<b>Grado de adherencia</b>	230		270		<b>0,007</b>
Superior a lo normal	160	69,6	162	60,0	
Normal	67	29,1	95	35,2	
Inferior a lo normal	3	1,3	13	4,8	

#### 4.1.4.2 Imagen corporal

La medición de la imagen corporal fue estimada mediante las dos subescalas del Eating Disorder Inventory (EDI): Insatisfacción Corporal (IC) y Obsesión por la Delgadez (OD). La evaluación final se muestra en la Tabla 19.

Tabla 19. *Distribución de la muestra total estudiada según las subescalas EDI-IC y EDI-OD*

Subescalas EDI	n(%)	Total	
		Media±DT	Mediana
EDI-IC	486	4,7±5,6	3,0
EDI-OD	486	4,5±4,5	3,0

Las puntuaciones de la medición de la imagen corporal mediante el EDI-IC y EDI-OD se muestran en las Figuras 9 Y 10. Mayores puntuaciones implican mayor insatisfacción corporal y mayor obsesión por la delgadez.

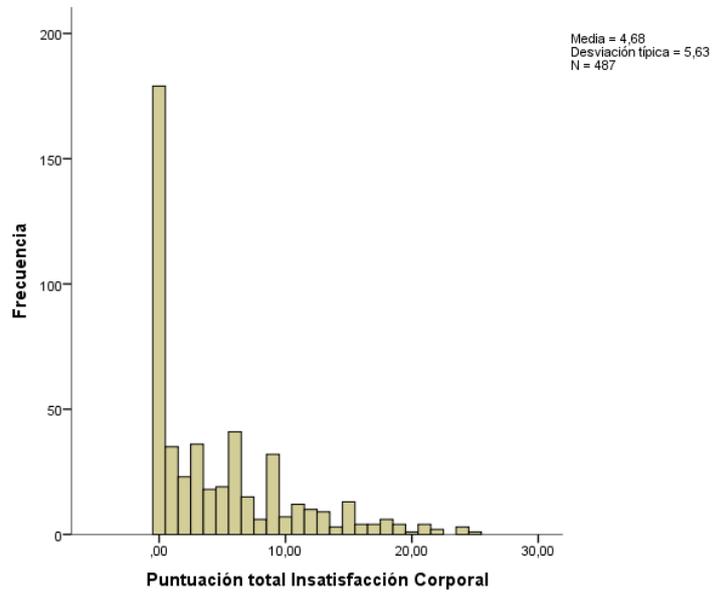


Figura 9. Distribución de pacientes según la puntuación total de la subescala insatisfacción corporal (EDI-IC)

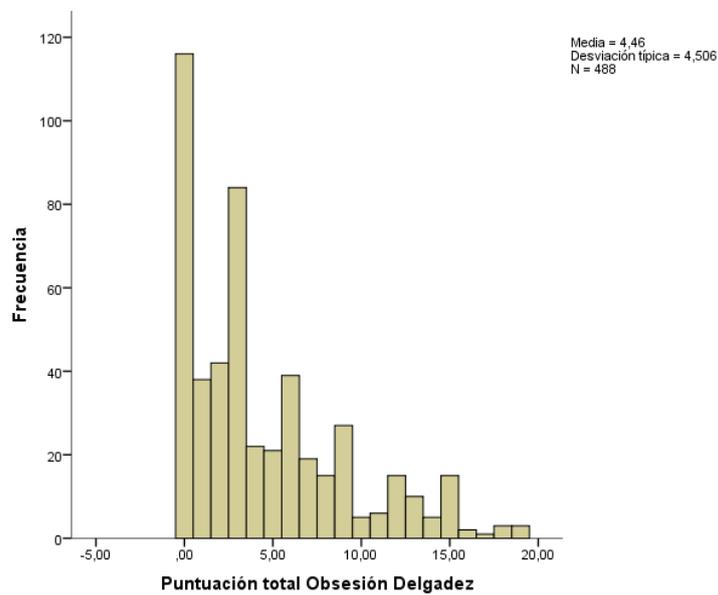


Figura 10. Distribución de pacientes según la puntuación total de la subescala obsesión por la delgadez (EDI-OD)

En la tabla 20, se detalla que existen diferencias estadísticamente significativas en la subescala EDI-IC, con una mayor insatisfacción corporal en el grupo de menor edad.

Tabla 20. *Distribución de la muestra estudiada según las subescalas EDI-IC y EDI-OD en los grupos de edad.*

Subescalas EDI	n(%)	40-64 años		n(%)	65 y más años		p
		Media±DT	Mediana		Media±DT	Mediana	
EDI-IC	230	5,8±5,9	4,5	256	3,6±5,1	0,0	<b>&lt;0,001</b>
EDI-OD	230	4,4±4,3	3,0	257	4,5±4,6	3,0	0,957

## 4.2. EVALUACIÓN DEL ESTADO NUTRICIONAL

### 4.2.1. PREVALENCIA Y RIESGO DE MALNUTRICIÓN SEGÚN EL MINI NUTRITIONAL ASSESSMENT (MNA)

La evaluación del estado nutricional fue estimada por medio del cuestionario Mini Nutritional Assessment. La distribución según las diferentes preguntas del cuestionario se muestra en la Tabla 21.

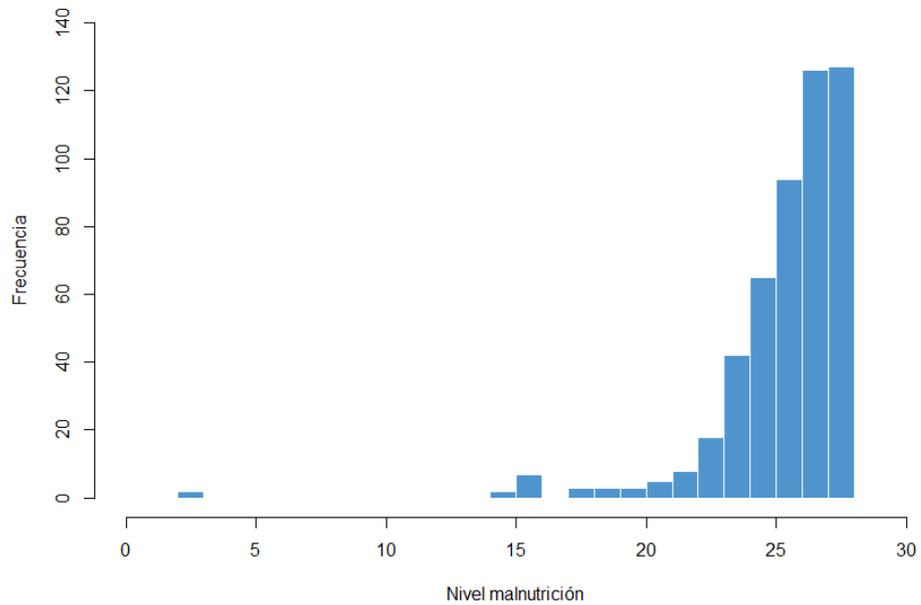
Tabla 21. *Distribución de la muestra estudiada según la evaluación del estado nutricional.*

<b>MNA (MINI NUTRITIONAL ASSESSMENT):</b>		<b>n</b>	<b>%</b>
<b>Evaluación del estado nutricional</b>			
<b>A. ¿Ha perdido el apetito? ¿Ha comido menos por falta de apetito, problemas digestivos, dificultades de masticación o deglución en los últimos 3 meses?</b>			
0 = ha comido mucho menos		8	1,6
1 = ha comido menos		39	7,8
2 = ha comido igual		456	90,7
<b>B. Pérdida reciente de peso (&lt;3 meses)</b>			
0 = pérdida de peso > 3 kg		31	6,2
1 = no lo sabe		0	0
2 = pérdida de peso entre 1 y 3 kg		60	11,9
3 = no ha habido pérdida de peso		412	81,9
<b>C. Movilidad</b>			
0 = de la cama al sillón		1	0,2
1 = autonomía en el interior		2	0,4
2 = sale del domicilio		500	99,4
<b>D. ¿Ha tenido una enfermedad aguda o situación de estrés psicológico en los últimos 3 meses?</b>			
0 = sí		116	23,2
2 = no		385	76,8
<b>E. Problemas neuropsicológicos</b>			
0 = demencia o depresión grave		1	0,2
1 = demencia moderada		2	0,4
2 = sin problemas psicológicos		496	99,4
<b>F. Índice de masa corporal (IMC) = peso en kg / (talla en m)<sup>2</sup></b>			
0 = IMC <19		0	0
1 = 19 ≤ IMC < 21		6	1,2
2 = 21 ≤ IMC < 23		28	5,6
3 = IMC ≥ 23		469	93,6
<b>G. ¿El paciente vive independiente en su domicilio?</b>			
1=sí		487	99,2
0=no		4	0,8
<b>H. ¿Toma más de 3 medicamentos al día?</b>			
0=sí		155	31,5
1=no		337	68,5
<b>I. ¿Úlceras o lesiones cutáneas?</b>			
0=sí		29	5,9
1=no		463	94,1
<b>J. ¿Cuántas comidas completas toma al día?</b>			
0 = 1 comida		10	2,0
1 = 2 comidas		51	10,3
2 = 3 comidas		436	87,7
<b>K. ¿Consume el paciente productos lácteos al menos una vez al día?</b>			
sí		473	95,2
no		24	4,8

## CONTINUACIÓN TABLA 21

<b>huevos o legumbres 1 o 2 veces a la semana?</b>		
si	494	99,4
no	3	0,6
<b>carne, pescado o aves, diariamente?</b>		
si	486	98
no	10	2,0
0,0= 0 o 1 síes	1	0,2
0,5= 2 síes	35	7,0
1,0= 3 síes	461	92,8
<b>L. ¿Consume frutas o verduras al menos 2 veces al día?</b>		
0 = no	56	11,3
1 = sí	441	88,7
<b>M. ¿Cuántos vasos de agua u otros líquidos toma al día? (agua, zumo, café, té, leche, vino, cerveza...)</b>		
0,0 = menos de 3 vasos	20	4,0
0,5 = de 3 a 5 vasos	188	37,9
1,0 = más de 5 vasos	288	58,1
<b>N. Forma de alimentarse</b>		
0 = necesita ayuda	2	0,4
1 = se alimenta solo con dificultad	2	0,4
2 = se alimenta solo sin dificultad	493	99,2
<b>O. ¿Se considera el paciente que está bien nutrido?</b>		
0 = malnutrición grave	2	0,4
1 = no lo sabe o malnutrición moderada	34	6,8
2 = sin problemas de nutrición	461	92,8
<b>P. En comparación con las personas de su edad, ¿cómo encuentra el paciente su estado de salud?</b>		
0,0 = peor	45	9,1
0,5 = no lo sabe	23	4,6
1,0 = igual	259	52,3
2,0 = mejor	168	33,9
<b>Q. Circunferencia braquial (CB en cm)</b>		
0,0 = CB < 21	0	0
0,5 = 21 ≤ CB ≤ 22	2	0,4
1,0 = CB > 22	500	99,6
<b>R. Circunferencia de la pantorrilla (CP en cm)</b>		
0 = CP < 31	46	9,2
1 = CP ≥ 31	453	90,8

La puntuación media en el cuestionario MNA, para el global de la muestra fue de 25,7±2,7 puntos, asociándose mayores puntuaciones a un menor riesgo de malnutrición (Figura 11). La prevalencia global de malnutrición o riesgo de malnutrición se encuentra en el 13,3% de los sujetos (Tabla 22).



Malnutrición (<17ptos)  
 Riesgo de malnutrición (17-23,5ptos)  
 Estado nutricional normal (≥24ptos)

Figura 11. Histograma de puntuación total del MNA (Mini Nutritional Assessment)

Tabla 22. Puntuación global del MNA y prevalencia de malnutrición y riesgo de malnutrición en la muestra total.

Evaluación del estado nutricional ( <i>Mini Nutritional Assessment</i> )	Total		
	n (%)	95%IC	Media±DT
<b>Puntuación global del MNA</b>	503		25,7±2,7
<b>Categorías del MNA</b>			
Malnutrición (<17ptos)	10(2,0)	0,7-3,3	2,0
Riesgo de malnutrición (17-23,5ptos)	57(11,3)	8,5-14,2	11,3
Estado nutricional normal (≥24ptos)	436(86,7)	83,7-89,7	86,7

Considerando la puntuación global del MNA existe un discreto incremento de dicho valor en la población de 65 y más años en comparación con los de 40-64 años. La malnutrición es mayor en el grupo más joven (3,4% vs 0,7%) (Tabla 23).

Tabla 23. Puntuación global del MNA y prevalencia de malnutrición y riesgo de malnutrición en los grupos de edad.

<i>Evaluación del estado nutricional (Mini Nutritional Assessment)</i>	40-64 años		65 y más años		p
	Media±DT	Mediana	Media±DT	Mediana	
<b>Puntuación global del MNA</b>	25,5±3,2	26,0	25,8±2,2	26,5	0,286
	<b>n</b>	<b>%</b>	<b>n</b>	<b>%</b>	<b>p</b>
<b>Categorías del MNA</b>	232		271		0,036
Malnutrición (<17ptos)	8	3,4	2	0,7	
Riesgo de malnutrición (17-23,5ptos)	21	9,1	36	13,3	
Estado nutricional normal (≥24ptos)	203	87,5	233	86,0	

#### 4.2.2. FACTORES ASOCIADOS AL RIESGO DE MALNUTRICIÓN

##### 4.2.2.1 Técnicas estadísticas

Se analizó el estado nutricional de los participantes de acuerdo a variables sociodemográficas, datos de comorbilidad, antropometría, y los resultados de adherencia a la dieta mediterránea y autopercepción de la imagen corporal.

En el análisis bivariado, la puntuación total obtenida por el MNA se correlaciona discreta pero significativamente con la puntuación del Charlson crudo, con la puntuación del cuestionario de Adherencia a la Dieta Mediterránea y con la puntuación obtenida en el cuestionario de insatisfacción corporal (Tabla 24). A su vez, se objetivan diferencias significativas en la puntuación del MNA y el hecho de vivir o no sólo (Tabla 25).

Tabla 24. Correlación de la puntuación MNA con diferentes variables.

<b>Malnutrición y riesgo de malnutrición</b>		
<b>Variable sociodemográfica</b>	<b>Rho de Spearman</b>	<b>p</b>
Edad (años)	0,01	0,812
<b>Comorbilidad</b>		
Charlson crudo	-0,10	<b>0,026</b>
Charlson ajustado	-0,02	0,617
TAS (Valor medio en el estudio)	0,05	0,282
TAD (Valor medio en el estudio)	0,06	0,205
Fageström total	0,05	0,715
<b>Adherencia a la dieta mediterránea</b>	0,10	<b>0,028</b>
<b>EDI (Insatisfacción corporal)</b>	0,09	<b>0,038</b>
<b>EDI (Obsesión por la delgadez)</b>	0,03	0,469
<b>Variables antropométricas</b>		
Peso (Kg)	0,05	0,253
IMC (kg/m <sup>2</sup> )	0,06	0,207
Perímetro de cintura (cm)	0,03	0,529
Perímetro de cadera (cm)	0,04	0,355
Índice cintura - cadera	0,03	0,536
Perímetro del brazo (cm)	0,03	0,481
Perímetro de pantorrilla (cm)	0,07	0,118
Pliegues cutáneos (mm)		
Tricipital	0,02	0,685
Bicipital	0,06	0,162
Subescapular	0,00	0,994
Suprailíaco	0,04	0,345
Perímetro muscular del brazo (cm)	0,04	0,394
Área muscular del brazo	0,04	0,394
Área grasa del brazo	0,03	0,509
Índice adiposo muscular	-0,00	0,945
Bioimpedancia		
Masa adiposa	-0,02	0,680
Masa líquida	0,02	0,612
Masa muscular	0,01	0,883
Masa ósea	0,04	0,318
Metabolismo basal calórico	0,05	0,269
Metabolismo en actividad	0,08	0,092

Del análisis de los datos se puede deducir que existe una correlación negativa entre la puntuación del MNA y el Charlson crudo (Rho Spearman= -0,10; p=0,026). Existe una correlación positiva entre el MNA y la Adherencia a la Dieta Mediterránea (Rho Spearman=0,10; p=0,028). A su vez, los pacientes con adherencia a la dieta mediterránea tienen mayor puntuación del MNA que los que no tienen dicha adherencia (26,1±1,84 vs 25,4±2,52)

También se puede deducir que, a mayor puntuación de la insatisfacción corporal, mayor puntuación del MNA (Rho Spearman=0,09; p=0,038) (Tabla 24). Los pacientes con Insatisfacción Corporal tienen una discreta mayor puntuación que los que no la tienen (26,0±1,9 vs 25,5±2,5) (Tabla 25).

Además, las personas que viven solas tienen significativamente menor puntuación del MNA que las que no viven solas (24,5±4,4 vs 25,8±2,49) (Tabla 25).

Tabla 25. Puntuación media del MNA según diferentes variables de interés

<b>Malnutrición y riesgo de malnutrición</b>	<b>Media±DT; n</b>	<b>P</b>
<b>Sexo</b>		0,663
Hombre	0,02±0,15; 207	
Mujer	0,02±0,13; 298	
<b>IMC</b>		0,228
Normopeso (IMC 20-24,9Kg/m <sup>2</sup> )	24,9±3,8; 83	
Sobrepeso (IMC 25-29, 9Kg/m <sup>2</sup> )	25,7±2,6; 229	
Obesidad (IMC >30Kg/ m <sup>2</sup> )	25,9±1,8; 189	
<b>Nivel de estudios</b>		0,421
Sin estudios	0,02±0,15; 91	
1º grado incompleto	0,02±0,13; 223	
1º grado (EGB)	0,02±0,14; 52	
2º grado	0,02±0,15; 88	
Universitario	0,0±0,0; 49	
<b>Estado civil</b>		0,910
Soltero	24,5±5,6; 23	
Casado/pareja	25,7±2,6; 386	
Viudo	25,6±2,14; 63	
Divorciado	26,3±1,41; 12	
Separado	26,4±1,43; 9	
<b>¿Vive sólo?</b>		<b>0,032</b>
Si	24,5±4,4; 37	
No	25,8±2,49; 456	
<b>Hábito tabáquico</b>		0,876
Fumador	25,8±2,2; 71	
Exfumador	25,6±2,6; 127	
No fumador	25,7±2,8; 305	
<b>Test de Fagerstrom</b>		0,284
Dependencia alta	25,9±1,5; 28	
Dependencia moderada	25,8±2,9; 31	
Dependencia baja	25,0±1,8; 6	
<b>HTA (TAS≥140mmHg y TAD≥90mmHg)</b>		0,724
Si	25,8±2,1; 152	

## CONTINUACIÓN TABLA 25

No	25,6±2,9; 353	
<b>Adherencia a la dieta mediterránea</b>		<b>0,007</b>
Si	26,1±1,84; 368	
No	25,4±2,52; 113	
<b>EDI (Insatisfacción corporal)</b>		<b>0,019</b>
Si	26,0±1,9; 250	
No	25,5±2,5; 237	
<b>EDI (Obsesión por la delgadez)</b>		<b>0,645</b>
Si	25,8±2,2; 292	
No	25,7±2,2; 196	

La puntuación total del MNA no se correlaciona ni con la edad, ni con la dependencia del tabaquismo, ni con la obsesión por la delgadez, ni con las variables antropométricas (peso, talla, IMC, perímetros, pliegues, áreas y bioimpedancia) (Tabla 24). Tampoco se asocia ni con el sexo ni con las categorías del IMC, ni con el nivel de estudios, ni con el estado civil, ni con la exposición al tabaco (Tabla 25).

Tras categorizar la puntuación del MNA en dos grupos (malnutrición (<17 puntos) y riesgo de malnutrición (17-23,5 puntos) en comparación con los que no tienen malnutrición (>24 puntos), se objetivan hallazgos en la misma tendencia (Tabla 26,27).

Tabla 26. Análisis bivariado para determinar variables asociadas a la presencia de malnutrición o riesgo de malnutrición según el MNA.

Malnutrición y riesgo de malnutrición	Sí Media±DT	No Media±DT	p	OR (95%IC)
<b>Variables sociodemográficas</b>				
<b>Edad</b> (años)	62,4±13,3	63,2±11,9	0,730	0,9(0,9-1,0)
	<b>n(%)</b>	<b>n(%)</b>		
<b>Sexo</b>			0,765	
Mujer	39 (57,4)	259 (59,3)		1
Hombre	29 (42,6)	178 (40,7)		1,1(0,6-1,8)
<b>¿Vive sólo?</b>			<b>0,020</b>	
No	56(84,8)	400(93,7)		1
Si	10(15,2)	27(6,3)		2,6(1,2-5,6)

CONTINUACIÓN TABLA 26				
	Media±DT	Media±DT		
<b>Comorbilidad</b>				
Charlson crudo	0,6±0,9	0,5±1,0	0,580	1,1(0,8-1,4)
Charlson ajustado	2,4±1,9	2,3±1,8	0,929	1,0(0,9-1,2)
TAS (Valor medio en el estudio)	123,3±14,8	128,8±14,8	<b>0,005</b>	0,9(0,9-0,9)
TAD (Valor medio en el estudio)	72,0±9,9	74,7±10,1	<b>0,041</b>	0,9(0,9-0,9)
Fageström total	6,9±2,2	7,1±2,6	0,781	0,9(0,7-1,3)
<b>Adherencia a la dieta total</b>	9,2±2,0	9,8±1,8	<b>0,030</b>	0,8(0,7-0,9)
	n(%)	n(%)		
<b>Adherencia a la dieta mediterránea</b>			0,142	
No	17(31,5)	96(22,5)		1
Si	37(68,5)	331(77,5)		1,6(0,8-2,9)
	Media±DT	Media±DT		
<b>Insatisfacción corporal</b>	3,7±5,8	4,8±5,6	0,136	0,9(0,9-1,0)
<b>Obsesión por la delgadez</b>	4,2±4,3	4,5±4,5	0,212	0,9(0,9-1,0)
<b>Variables antropométricas</b>				
Peso (Kg)	72,3±17,7	74,9±13,2	0,150	0,9(0,9-1,0)
Talla (cm)	159,4±10,2	160,3±9,4	0,481	0,9(0,9-1,0)
IMC (kg/m <sup>2</sup> )	28,3±6,3	29,1±4,4	0,172	0,9(0,9-1,0)
Perímetro de cintura	91,0±14,3	93,9±11,8	<b>0,048</b>	0,9(0,9-1,0)
Perímetro de cadera	101,7±12,6	102,5±9,2	0,180	0,9(0,9-1,0)
Índice cintura - cadera	0,89±0,09	0,92±0,09	0,057	0,1(0,0-1,2)
Perímetro del brazo	29,4±4,2	30,1±3,3	0,106	0,9(0,9-1,0)
Perímetro de pantorrilla	34,2±4,2	34,8±3,1	<b>0,043</b>	0,9(0,9-1,0)
<b>Pliegues cutáneos</b>				
Tricipital	15,7±5,7	15,9±4,8	0,610	0,9(0,9-1,0)
Bicipital	13,6±7,2	14,5±5,4	0,052	0,9(0,9-1,0)
Subescapular	18,8±6,6	19,8±5,2	0,056	0,9(0,9-1,0)
Suprailíaco	18,4±7,0	18,8±6,4	0,506	0,9(0,9-1,0)
P. muscular del brazo	24,5±3,4	25,2±3,1	0,169	0,9(0,9-1,0)
Área muscular del brazo	7,9±2,2	8,3±2,0	0,169	0,9(0,8-1,0)
Área grasa del brazo	3,5±1,8	3,6±1,2	0,361	0,9(0,8-1,2)
Índice adiposo muscular	0,4±0,16	0,4±0,16	0,789	1,1(0,2-5,3)
<b>Bioimpedancia</b>				
Masa adiposa	32,8±8,4	33,8±6,8	0,537	0,9(0,9-1,0)
Masa líquida	48,7±5,5	48,3±4,9	0,677	1,0(0,9-1,1)
Masa muscular	32,3±5,8	31,9±4,9	0,795	1,0(0,9-1,1)
Masa ósea	6,7±2,6	7,0±2,3	0,211	0,9(0,8-1,1)
Metabolismo basal calórico	1383,8±253,5	1429,5±224,4	0,074	0,9(0,9-1,0)
Metabolismo en actividad	2007,5±463,0	2071,1±405,8	0,088	1,0(0,9-1,0)

Tras realizar un análisis multivariado de regresión logística teniendo en cuenta la edad, el vivir sólo, el sexo, el score de comorbilidad de Charlson, la adherencia a la dieta mediterránea y el perímetro de cintura y de pantorrilla (Tabla 27) se objetiva que las variables asociadas de manera independiente al riesgo de malnutrición son: el vivir sólo (OR=2,92); una mayor edad (OR=1,04), y un perímetro de cintura menor (OR=0,96). Encontrándose la adherencia a la dieta mediterránea en el límite de la significación estadística (OR=0,53; p=0,082).

Tabla 27. Análisis multivariado de regresión logística para determinar variables asociadas a la presencia de malnutrición o riesgo de malnutrición según el MNA.

Malnutrición y riesgo de malnutrición	B	EE	p	OR	95% IC (OR)	
¿Vive solo?	1.073	.466	<b>.021</b>	2.925	1.173	7.298
Edad	.036	.016	<b>.029</b>	1.036	1.004	1.070
Sexo	-.408	.370	.271	.665	.322	1.374
Charlson crudo	.067	.175	.701	1.070	.759	1.507
Adherencia a la dieta mediterránea	-.637	.366	.082	.529	.258	1.084
Perímetro de pantorrilla	-.016	.068	.817	.984	.862	1.125
Perímetro de cintura	-.041	.020	<b>.045</b>	.960	.923	.999
Constante	.345	2.366	.884	1.411		

#### 4.2.2.2 Técnicas de inteligencia artificial

Los resultados que se analizan a continuación han utilizado un total de 38 variables para tratar de mostrar las diferencias entre la existencia o no de malnutrición. Una vez preprocesados, los datos se han estandarizado de tal manera que cada variable se ha ajustado a los ejemplos, para que tengan una media igual a cero y una desviación típica igual a 1.

Para el análisis de los datos se han utilizado tres técnicas diferentes de ML en la tarea de verificar los resultados obtenidos, para tratar de identificar el modelo que ofrece mejores resultados en general, y la técnica con menor grado de variabilidad, lo que permite estudiar la importancia del resultado de las variables seleccionadas. Además, se ha realizado un proceso de selección de características, para intentar reducir lo máximo

posible el número de variables utilizadas y que los resultados sean similares, sin diferencias estadísticas, si no mejores, para las obtenidas utilizando todas las variables.

A continuación, en la Figura 12 se muestran los “boxplots” con la mediana de resultados tras 50 repeticiones de “10-fold cross-validation”. En cuanto a las métricas de rendimiento: en a) se expone el RMSE; y en b) se muestra el coeficiente de determinación, en este caso la “Rho de Spearman”.

Se observa que los modelos presentan unos resultados muy variables con grandes oscilaciones debido a la presencia de *outliers*. Esto lleva a pensar que es un problema complejo y que existe poca información en los datos o con insuficiente capacidad discriminativa.

GLMNET es el mejor modelo, con un valor ponderado cercano a 0,8, pero aun así no parece un modelo suficientemente estable para aplicación clínica.

Los test estadísticos indican que no hay diferencias significativas entre los modelos debido a la gran variabilidad de los resultados.

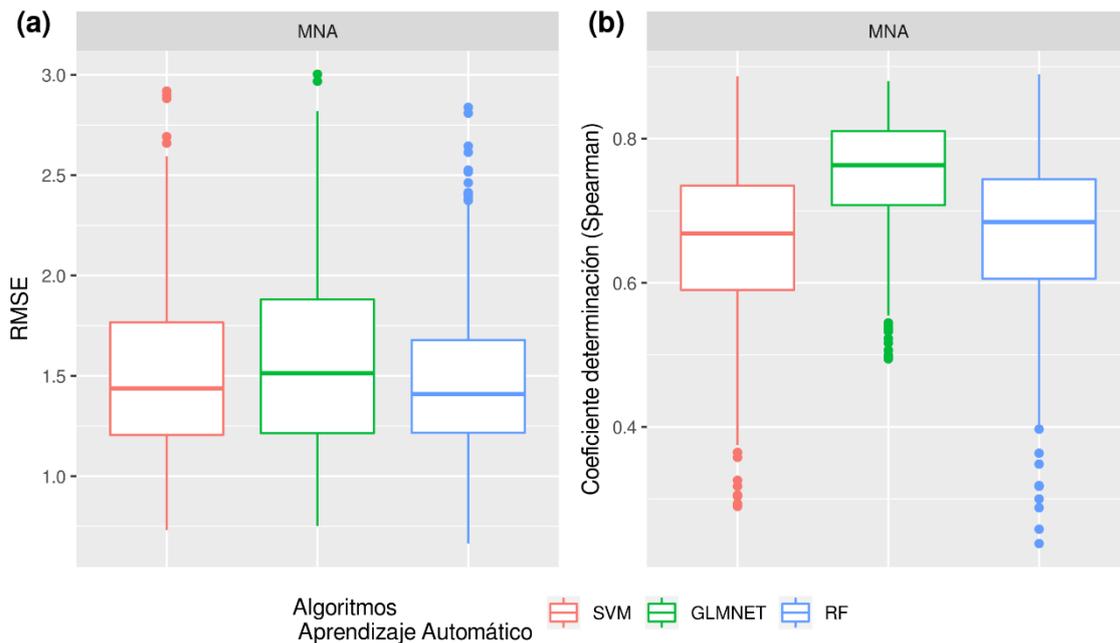


Figura 12. “Boxplots” de resultados de los algoritmos de aprendizaje automático

### 4.3. ADHERENCIA A LA DIETA MEDITERRÁNEA

#### 4.3.1. GRADO DE ADHERENCIA A LA DIETA MEDITERRÁNEA

La distribución de las respuestas al cuestionario de adherencia a la DMed se muestra en la Tabla 28.

Tabla 28. *Distribución de la muestra estudiada según la adherencia a la dieta mediterránea*

Adherencia a la dieta mediterránea	n	%
<b>1. ¿Usa usted el aceite de oliva principalmente para cocinar?</b>		
Si (1 punto)	459	95,4
No	22	4,3
<b>2. ¿Cuánto aceite de oliva consume en total al día (incluyendo el usado para freír, el de las comidas fuera de casa, las ensaladas...)?</b>		
Menos de dos cucharadas	46	9,6
Dos cucharadas o más (1 punto)	435	90,4
<b>3. ¿Cuántas raciones de verdura u hortalizas consume al día (las guarniciones o acompañamientos contabilizan como ½ ración)?</b>		
Menos de dos al día	190	39,5
Dos o más al día (al menos una de ellas en ensaladas o cruda= 1 punto)	291	60,5
<b>4. ¿Cuántas piezas de fruta (incluyendo zumo natural) consume al día?</b>		
Menos de tres al día	208	43,2
Tres o más al día (1 punto)	273	56,8
<b>5. ¿Cuántas raciones de carnes rojas, hamburguesas, salchichas o embutidos consume al día (una ración equivale a 100-150 gr.)?</b>		
Más de una al día	27	5,6
Menos de una al día (1 punto)	454	94,4
<b>6. ¿Cuántas raciones de mantequilla, margarina o nata consume al día (una porción individual equivale a 12 gr.)?</b>		
Más de una al día	51	10,6
Menos de una al día (1 punto)	429	89,4
<b>7. ¿Cuántas bebidas carbonatadas y/o azucaradas (refrescos, colas, tónicas, bitter) consume al día?</b>		
Más de una al día	64	13,3
Menos de una al día (1 punto)	417	86,7
<b>8. ¿Bebe vino? ¿Cuánto consume a la semana?</b>		
Menos de tres vasos por semana	304	63,2
Tres o más vasos por semana (1 punto)	177	36,8

## CONTINUACIÓN TABLA 28

<b>9. ¿Cuántas raciones de legumbres consume a la semana (una ración o plato equivale a 150 gr)?</b>		
Menos de tres por semana	166	34,5
Tres o más por semana (1 punto)	315	65,5
<b>10. ¿Cuántas raciones de pescado o mariscos consume a la semana (un plato, pieza o ración equivale a 100-150 gr de pescado ó 4-5 piezas de marisco)?</b>		
Menos de tres por semana	212	44,1
Tres o más por semana (1 punto)	269	55,9
<b>11. ¿Cuántas veces consume repostería comercial (no casera) como galletas, flanes, dulces o pasteles a la semana?</b>		
Menos de tres por semana (1 punto)	427	88,8
Tres o más por semana	54	11,2
<b>12. ¿Cuántas veces consume frutos secos a la semana (una ración equivale a 30 gr)?</b>		
Ninguna	259	54,0
Una o más por semana (1 punto)	221	46,0
<b>13. ¿Consume preferentemente carne de pollo, pavo o conejo en vez de ternera, cerdo, hamburguesas o salchichas (carne de pollo: una pieza o ración equivale a 100- 150 gr)?</b>		
No	136	28,3
Si (1 punto)	345	71,7
<b>14. ¿Cuántas veces a la semana consume los vegetales cocinados, la pasta, el arroz u otros platos aderezados con una salsa de tomate, ajo, cebolla o puerro elaborada a fuego lento con aceite de oliva (sofrito)?</b>		
Menos de dos por semana	316	65,7
Dos o más por semana (1 punto)	165	34,3

La puntuación media en el cuestionario de adherencia a la dieta mediterránea fue de  $9,7 \pm 1,8$  puntos, con una mediana de 10 puntos y un rango de 2 a 14, indicando mayor puntuación mejor adherencia a dicha dieta (Figura 13). La evaluación final de la adherencia a la dieta mediterránea, según sea baja ( $<9$  puntos) o alta ( $\geq 9$  puntos), se muestra en la Tabla 29 y la Figura 14, donde se objetiva que el 23,4% (95% IC=19,5-27,3) de los pacientes presentan baja adherencia a la dieta mediterránea.

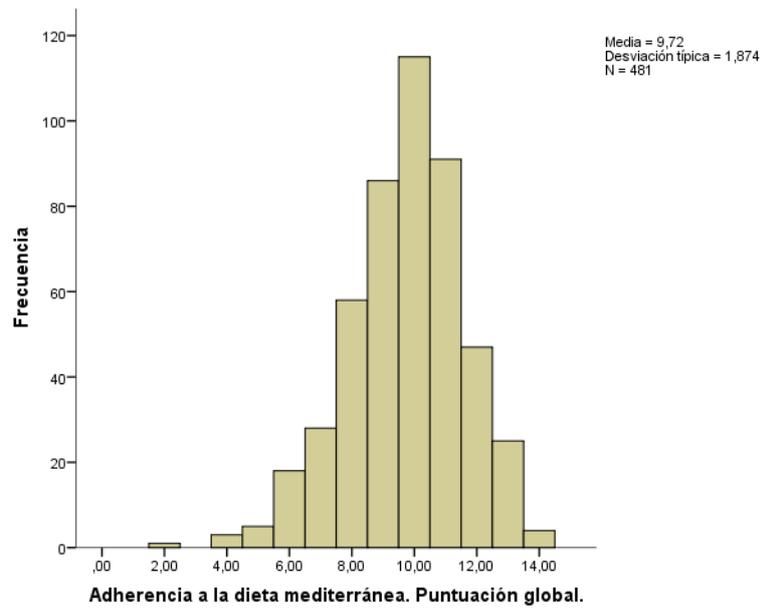


Figura 13. Distribución de las puntuaciones del cuestionario de adherencia a la dieta mediterránea.

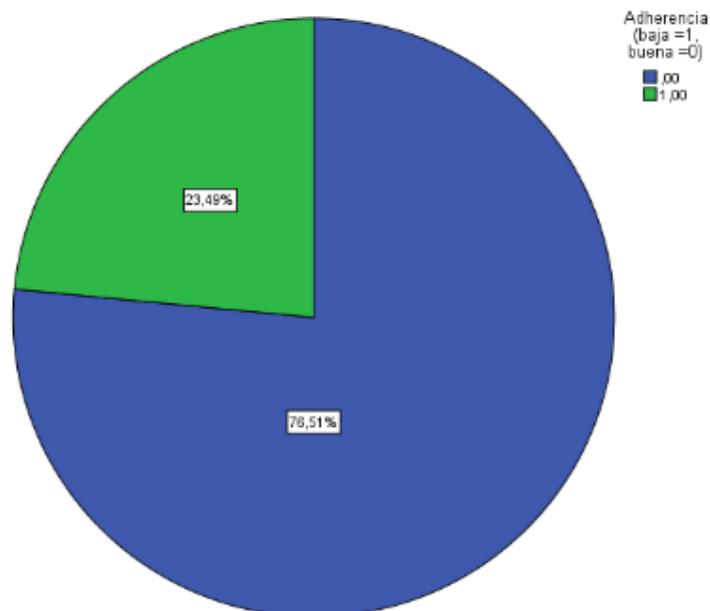


Figura 14. Distribución de pacientes según la adherencia a la dieta mediterránea

Tabla 29. Puntuación global y categórica según la adherencia a la dieta mediterránea en la muestra total

Adherencia a la dieta mediterránea	n (%)	Total de la muestra		
		95%IC	Media±DT	Mediana
<b>Puntuación global</b>	479		9,7±1,8	10,0
<b>Grado de adherencia</b>				
Baja adherencia (<9ptos)	112(23,4)	19,5-27,3		
Buena adherencia (≥9ptos)	367(76,6)	72,7-80,5		

Considerando la puntuación global del cuestionario de la DM, se objetiva una discreta mayor puntuación a la adherencia a la dieta mediterránea en la gente de mayor edad. Tras categorizar los valores en baja (<9ptos) y buena adherencia (≥9ptos), la diferencia no es significativa entre los grupos de edad (Tabla 30). Así, un 74,5% de las personas entre 40 y 64 años presentan buena adherencia, siendo este porcentaje del 78,2% entre los de 65 años o más (p=0,342).

Tabla 30. Puntuación global y categórica según la adherencia a la dieta mediterránea en los grupos de edad

Adherencia a la dieta mediterránea	40-64 años		65 y más años		P
	Media±DT	Mediana	Media±DT	Mediana	
<b>Puntuación global</b>	9,6±1,9	10,0	9,9±1,8	10,0	0,108
	<b>n</b>	<b>%</b>	<b>n</b>	<b>%</b>	<b>p</b>
<b>Grado de adherencia</b>	208		271		0,342
Baja adherencia (<9ptos)	53	25,5	59	21,8	
Buena adherencia (≥9ptos)	155	74,5	212	78,2	

#### 4.3.2. FACTORES ASOCIADOS A LA ADHERENCIA A LA DIETA MEDITERRÁNEA

##### 4.3.2.1 Técnicas estadísticas

En las Tablas 31 y 32, se analiza la asociación de la puntuación total de adherencia a la DMed con diferentes variables.

Con respecto a las variables cuantitativas, existe una correlación discreta pero significativa entre una mayor adherencia a la dieta mediterránea y una menor comorbilidad medida por el score de Charlson, una mayor dependencia a la nicotina según el test de Fageström, mejor estado nutricional según el MNA, mayor satisfacción corporal y mayor obsesión por la delgadez. En cuanto a los parámetros antropométricos, se objetivó una correlación negativa con el peso, IMC, perímetro de cintura, índice cintura-cadera, perímetro del brazo, perímetro de pantorrilla, pliegue bicipital, pliegue subescapular, perímetro muscular del brazo y área muscular del brazo. Con respecto a los resultados de la bioimpedancia, se detectó una correlación significativa con valores más bajos de masa ósea y metabolismo basal calórico (Tabla 31).

Tabla 31. *Correlación de la puntuación del cuestionario de adherencia a la dieta mediterránea con diferentes variables*

<b>Adherencia a la dieta mediterránea</b>		
<b>Variable sociodemográfica</b>	<b>Rho de Spearman</b>	<b>p</b>
Edad (años)	0,7	0,133
<b>Comorbilidad</b>		
Charlson crudo	-0,11	<b>0,016</b>
Charlson ajustado	0,02	0,617
TAS (Valor medio en el estudio)	-0,06	0,190
TAD (Valor medio en el estudio)	0,00	0,996
Fageström total	0,44	<b>0,001</b>
<b>MNA: Evaluación del estado nutricional</b>	0,10	<b>0,028</b>
<b>EDI (Insatisfacción corporal)</b>	-0,10	<b>0,031</b>
<b>EDI (Obsesión por la delgadez)</b>	0,09	<b>0,049</b>
<b>Variables antropométricas</b>		
Peso (Kg)	-0,14	<b>0,002</b>
IMC (kg/m <sup>2</sup> )	-0,10	<b>0,024</b>
Perímetro de cintura (cm)	-0,15	<b>0,001</b>
Perímetro de cadera (cm)	-0,09	0,053
Índice cintura - cadera	-0,13	<b>0,006</b>

CONTINUACIÓN TABLA 31

Perímetro del brazo (cm)	-0,14	<b>0,001</b>
Perímetro de pantorrilla (cm)	-0,11	<b>0,016</b>
Pliegues cutáneos (mm)		
Tricipital	-0,04	0,386
Bicipital	-0,09	<b>0,046</b>
Subescapular	-0,11	<b>0,015</b>
Suprailíaco	-0,06	0,177
Perímetro muscular del brazo (cm)	-0,14	<b>0,002</b>
Área muscular del brazo	-0,14	<b>0,002</b>
Área grasa del brazo	-0,06	0,210
Índice adiposo muscular	0,004	0,934
<b>Bioimpedancia</b>		
Masa adiposa	-0,05	0,238
Masa líquida	0,02	0,261
Masa muscular	-0,08	0,094
Masa ósea	-0,11	<b>0,018</b>
Metabolismo basal calórico	-0,15	<b>0,001</b>
Metabolismo en actividad	-0,09	0,060

Además, se muestra también de forma significativa que las mujeres tienen mayor adherencia a la dieta mediterránea que los hombres ( $9,9 \pm 1,7$  vs  $9,5 \pm 2,1$ ). Asimismo, las personas con una dependencia baja a la nicotina, según el test de Fageström, presentan menor adherencia a la dieta mediterránea ( $7,0 \pm 0,9$ ) que las que presentan una adherencia moderada ( $9,2 \pm 1,8$ ) o alta ( $9,9 \pm 1,9$ ) ( $p=0,002$ ). A su vez, los sujetos en riesgo de malnutrición, según los resultados del MNA, reflejan puntuaciones de adherencia más bajas ( $9,2$  vs.  $9,8$ ;  $p=0,030$ ). Finalmente, aquellos que tienen obsesión por la delgadez, en comparación con los que no la tienen, poseen mayor puntuación en la adherencia a la dieta mediterránea ( $9,8$  vs  $9,5$ ;  $p=0,035$ ) (Tabla 31).

Tabla 32. Puntuación media del cuestionario de adherencia a la dieta mediterránea según diferentes variables de interés

<b>Adherencia a la dieta mediterránea</b>		
	<b>Media±DT; n</b>	<b>P</b>
<b>Sexo</b>		<b>0,027</b>
Hombre	9,5±2,1; 191	
Mujer	9,9±1,7; 290	
<b>IMC</b>		0,152
Normopeso (IMC 20-24,9Kg/m <sup>2</sup> )	9,9±1,9; 75	
Sobrepeso (IMC 25-29, 9Kg/m <sup>2</sup> )	9,8±1,7; 217	
Obesidad (IMC >30Kg/ m <sup>2</sup> )	9,6±1,9; 185	
<b>Nivel de estudios</b>		0,098
Sin estudios	9,4±1,9; 91	
1º grado incompleto	9,9±1,7; 220	
1º grado (EGB)	9,2±2,3; 48	
2º grado	9,6±1,9; 75	
Universitario	10,3±1,6; 45	
<b>Estado civil</b>		0,445
Soltero	9,8±1,5; 18	
Casado/pareja	9,8±1,8; 367	
Viudo	9,5±1,8; 63	
Divorciado	9,8±2,2; 12	
Separado	9,2±1,9; 9	
<b>¿Vive sólo?</b>		0,215
Si	9,4±1,9; 34	
No	9,8±1,8; 435	
<b>Hábito tabáquico</b>		0,226
Fumador	9,3±1,9; 64	
Exfumador	9,7±1,9; 118	
No fumador	9,8±1,8; 297	
<b>Test de Fagerstrom</b>		<b>0,002</b>
Dependencia alta	9,9±1,9; 25	
Dependencia moderada	9,2±1,8; 27	
Dependencia baja	7,0±0,9; 6	
<b>HTA (TAS≥140mmHg y TAD≥90mmHg)</b>		0,501
Si	9,6±1,9; 149	
No	9,7±1,8; 332	
<b>Malnutrición y riesgo de malnutrición</b>		<b>0,030</b>
Si	9,2±2,0; 54	
No	9,8±1,8; 427	
<b>EDI-IC (Insatisfacción corporal)</b>		0,166
Si	9,6±1,8; 240	
No	9,8±1,9; 225	
<b>EDI-OD (Obsesión por la delgadez)</b>		<b>0,035</b>
Si	9,8±1,7; 278	
No	9,5±2,04; 188	

Si se categoriza la puntuación de la adherencia a la dieta mediterránea en: buena adherencia ( $\geq 9$ ptos) y baja adherencia ( $< 9$ ptos); se observa que aquellos con buena adherencia a la DM presentan un peso significativamente más bajo (73,3 vs. 77,4;  $p=0,011$ ), mayor dependencia a la nicotina (Test de Fagérstrom 7,9 vs. 5,5;  $p=0,001$ ), mayor puntuación del MNA (26,1 vs. 25,4;  $p=0,007$ ) y mayor obsesión por la delgadez (4,8 vs. 3,8;  $p=0,014$ ). Con respecto a las características antropométricas, los sujetos con buena adherencia a la DM tienen valores significativamente más bajos de perímetro de cintura, índice cintura-cadera, perímetro de brazo, pliegue subescapular y perímetro y área muscular del brazo. En cuanto a los resultados de bioimpedancia, presentan valores también más bajos de masa ósea y metabolismo basal calórico (Tabla 32).

Tabla 33. *Análisis bivariado para determinar variables cuantitativas asociadas a la buena o baja adherencia a la dieta mediterránea*

Adherencia a la dieta mediterránea	Sí (Buena) Media±DT	No (Baja) Media±DT	p	OR (95% IC)
<b>VARIABLES SOCIODEMOGRÁFICAS</b>				
Edad	64,2±11,5	62,8±13,2	0,311	0,9(0,9-1,0)
Peso	73,3±12,8	77,4±14,7	<b>0,011</b>	1,0(1,0-1,0)
Talla	159,2±8,9	161,5±10,6	0,062	1,0(1,0-1,0)
IMC	28,9±4,3	29,6±4,7	0,102	1,0(0,9-1,1)
<b>COMORBILIDAD</b>				
Charlson crudo	0,5±0,9	0,7±1,1	0,058	1,2(0,9-1,5)
Charlson ajustado	2,5±1,8	2,4±1,9	0,961	0,9(0,9-1,1)
TAS (Valor medio en el estudio)	127,9±14,6	130,9±15,9	0,068	1,0(0,9-1,0)
TAD (Valor medio en el estudio)	74,5±10,1	73,9±10,8	0,607	0,9(0,9-1,0)
Fageström total	7,9±1,9	5,5±2,9	<b>0,001</b>	0,6(0,5-0,8)
<b>MNA total</b>	26,1±1,8	25,4±2,5	<b>0,007</b>	0,9(0,8-0,9)
<b>Insatisfacción corporal</b>	4,6±5,7	5,1±5,5	0,328	1,0(0,9-1,1)
<b>Obsesión por la delgadez</b>	4,8±4,6	3,8±4,6	<b>0,014</b>	0,9(0,9-1,0)
<b>VARIABLES ANTROPOMÉTRICAS</b>				
Perímetro de cintura	92,8±12,2	95,8±10,9	<b>0,009</b>	1,0(1,0-1,0)
Perímetro de cadera	102,1±9,1	103,3±9,7	0,161	1,0(0,9-1,0)
Índice cintura - cadera	0,9±0,1	0,9±0,1	<b>0,035</b>	0,1(0,1-0,9)
Perímetro del brazo	29,8±3,2	30,6±3,5	<b>0,041</b>	1,1(1,0-1,1)
Perímetro de pantorrilla	34,4±3,1	35,1±3,3	0,064	1,1(1,0-1,1)
<b>PLIEGUES CUTÁNEOS</b>				
Tricipital	15,7±4,7	15,7±5,3	0,872	0,9(0,9-1,0)
Bicipital	14,1±5,5	15,0±5,2	0,091	1,0(0,9-1,0)
Subescapular	19,2±5,3	20,8±5,4	<b>0,012</b>	1,0(1,0-1,0)
Suprailíaco	18,6±6,4	18,9±6,8	0,756	1,0(0,9-1,0)
P, muscular del brazo	24,9±3,0	25,6±3,3	<b>0,023</b>	1,1(1,0-1,1)
Área muscular del brazo	8,1±1,9	8,6±2,2	<b>0,023</b>	1,1(1,0-1,2)
Área grasa del brazo	3,5±1,2	3,6±1,4	0,625	1,1(0,9-1,2)
Índice adiposo muscular	0,4±0,2	0,4±0,2	0,523	0,6(0,2-2,2)
<b>BIOIMPEDANCIA</b>				
Masa adiposa	33,8±6,8	34,3±7,7	0,524	1,0(0,9-1,0)
Masa líquida	48,3±4,8	47,9±5,6	0,549	0,9(0,9-1,0)
Masa muscular	31,5±4,7	32,6±5,5	0,055	1,0(0,9-1,1)
Masa ósea	6,7±2,1	7,3±2,8	<b>0,011</b>	1,1(1,0-1,2)
Metabolismo basal calórico	1395,5±203,7	1476,7±270,1	<b>0,001</b>	1,0(1,0-1,0)
Metabolismo en actividad	2027,0±369,3	2111,2±462,8	0,290	1,0(0,9-1,0)

En cuanto a las características demográficas, no existen diferencias significativas entre las personas con baja o alta adherencia a la DM en cuanto a su nivel de estudios o estado civil. Existe un porcentaje de mujeres significativamente mayor entre aquellos con alta adherencia a la DM (79,7% vs. 52,2%;  $p=0,045$ ). Hay un mayor porcentaje de personas que viven solas entre los que presentan baja adherencia a la DM (6,1% vs. 11,1%), aunque las diferencias no alcanzan la significación estadística (Tabla 33).

Existe un menor porcentaje de fumadores entre aquellos con alta adherencia a la DM (11,4% vs. 19,6%;  $p=0,056$ ). Sin embargo, hay un mayor porcentaje de fumadores con dependencia alta a la nicotina entre los de mayor adherencia a la DM (52,6% vs. 25,0%;  $p=0,001$ ). La obsesión por la delgadez según el test EDI es más frecuente entre los de alta adherencia a la DM (62,3% vs. 50,9%,  $p=0,035$ ) (Tabla 34).

Tabla 34. *Análisis bivariado para determinar variables cualitativas asociadas a la buena o baja adherencia a la dieta mediterránea*

<b>Adherencia a la dieta mediterránea</b>	<b>Sí (Buena) Media±DT</b>	<b>No (Baja) Media±DT</b>	<b>p</b>	<b>OR</b>
<b>Edad agrupada</b>			0,342	
< 65 años	155(42,2)	53(47,3)		1
≥ 65 años	212(57,8)	59(52,7)		0,8(0,5-1,2)
<b>Sexo</b>			<b>0,045</b>	
Mujer	231(79,7)	59(52,2)		1
Hombre	137(20,3)	54(28,3)		1,5(1,0-2,4)
<b>IMC</b>			0,333	
Normopeso (IMC 20-24,9Kg/m <sup>2</sup> )	60 (16,4)	15 (13,4)		1
Sobrepeso (IMC 25-29, 9Kg/m <sup>2</sup> )	170 (46,6)	47 (42,0)		1,1(0,6-2,2)
Obesidad (IMC >30Kg/ m <sup>2</sup> )	135 (37,0)	50 (44,6)		1,5(0,8-2,9)
<b>Nivel estudios</b>			0,074	
Sin estudios	63(17,1)	28(25,2)		1
1° grado incompleto	178(48,4)	42(37,8)		0,5(0,3-0,7)
1° grado EGB	32(8,7)	16(14,4)		1,1(0,3-0,9)
2° grado	58(15,8)	17(15,3)		0,6(0,5-2,4)
Universidad	37(10,1)	8(7,2)		0,5(0,2-1,1)
<b>Estado civil</b>			0,517	
Soltero	13(3,6)	5(4,6)		1
Casado/pareja	285(78,9)	82(75,2)		0,7(0,3-2,4)
Viudo	48(13,3)	15(13,8)		0,8(0,2-2,9)
Divorciado	9(2,5)	3(2,8)		0,9(0,1-4,5)
Separado	6(1,7)	3(2,8)		1,3(0,2-7,3)
<b>¿Vive sólo?</b>			0,078	
No	339(93,9)	96(88,9)		1
Si	22(6,1)	12(11,1)		1,9(0,9-3,9)

CONTINUACIÓN TABLA 34

<b>Hábito tabáquico</b>			0,056	
No fumador	236(64,3)	61(54,5)		1
Exfumador	89(24,3)	29(25,9)		1,3(0,7-2,1)
Fumador	42(11,4)	22(19,6)		2,0(1,1-3,6)
<b>Test de Fagerstrom</b>			<b>0,001</b>	
Dependencia alta	20(52,6)	5(25,0)		1
Dep, moderada	18(47,4)	9(45,0)		2,0(0,6-7,1)
Dependencia baja	0	6(30,0)		6,5(0-NA)
<b>HTA</b>			0,163	
No	260(70,7)	72(63,7)		1
Si	108(29,3)	41(36,3)		1,4(0,9-2,1)
<b>Malnutrición o riesgo</b>			0,142	
No	331(89,9)	96(85,0)		1
Si	37(10,1)	17(15,0)		1,6(0,8-2,9)
<b>EDI (Insatisfacción corporal)</b>			0,248	
No	178(49,9)	47(43,5)		1
Si	179(50,1)	61(56,5)		1,3(0,8-1,9)
<b>EDI (Obsesión por la delgadez)</b>			<b>0,035</b>	
No	135(37,7)	53(49,1)		1
Si	223(62,3)	55(50,9)		0,6(0,4-0,9)

Tras realizar un análisis multivariado de regresión logística teniendo en cuenta la edad, el IMC, el hábito tabáquico, la puntuación del cuestionario MNA, la puntuación en obsesión por la delgadez, parámetros antropométricos y los resultados de la bioimpedancia; se objetiva que las variables asociadas de manera independiente a una baja adherencia a la dieta mediterránea son: ser fumador (OR=2,0); un mayor riesgo de malnutrición según el cuestionario MNA (OR=0,8); una menor obsesión por la delgadez (OR=0,9); valores más elevados del pliegue subescapular (OR=1,78) y masa muscular (OR=1,2); valores más bajos de masa ósea (OR=0,3); y un mayor metabolismo basal calórico (OR=1,0) (Tabla 35)

*Tabla 35. Análisis multivariante de regresión logística para determinar variables asociadas de forma independiente a una baja adherencia a la dieta mediterránea.*

<b>Adherencia a la dieta mediterránea</b>	<b>B</b>	<b>EE</b>	<b>P</b>	<b>OR</b>	<b>95% IC (OR)</b>	
Edad	-0.030	0.018	0.088	0.970	0.937	1.005
IMC	-0.005	0.048	0.921	0.995	0.907	1.093
Fumador	0.697	0.352	0.048	2.008	1.007	4.005
Cuestionario MNA	-0.222	0.059	<0.001	0.801	0.713	0.899
Puntuación Obsesión Delgadez	-0.078	0.030	0.009	0.925	0.872	0.980
Pliegue subescapular	0.580	0.288	0.044	1.785	1.015	3.141
Masa muscular	0.172	0.077	0.026	1.187	1.020	1.381
Masa ósea	-1.028	0.375	0.006	0.358	0.172	0.746
Metabolismo basal calórico	0.008	0.003	0.003	1.008	1.003	1.014
Constante	-4.412	3.026	0.145			

#### 4.3.2.2 Técnicas de inteligencia artificial

Los resultados y los modelos ML que se muestran a continuación, han utilizado un total de 38 variables para tratar de caracterizar las diferencias que existen entre la adhesión alta y baja a la dieta mediterránea. Al igual que se ha hecho con los datos de la evaluación nutricional; una vez preprocesados, se han estandarizado de tal manera que cada variable tiene una media igual a cero y una desviación típica igual a 1.

Se han utilizado las cuatro técnicas de ML descritas anteriormente: Random Forest, SVM, Elastic Net (GLMNET) y KNN, tratando de identificar el modelo que aporta mejores resultados. Al igual que con el análisis nutricional se ha realizado un proceso de selección de características para intentar reducir lo máximo posible el número de variables utilizadas.

El resultado agregado promedio de la ejecución de los diez experimentos de validación cruzada de 10 iteraciones que no tienen en cuenta el proceso de búsqueda interno completo de los mejores parámetros de cada técnica, sino el resultado de la aplicación de los mejores encontrados, se puede ver en Figura 15.

En la separación de los ejemplos en los diferentes subconjuntos de acuerdo con la técnica de validación cruzada utilizada, los ejemplos se han estratificado de tal manera que en los conjuntos de validación siempre hay un número proporcional de casos de la

clase minoritaria para evitar sesgos en el resultados A continuación, se muestran los resultados de los subconjuntos de tamaños de variables (4,16,32), identificados mediante la aplicación de una prueba “t” después de una selección de características de filtro de aproximación, para cuantificar la correlación entre cada una de las variables y las etiquetas que pertenecen a cada clase de ejemplos. Además, los resultados finales obtenidos al utilizar todo el conjunto de variables, como puede verse en la Figura 15, es una clara tendencia creciente que, a medida que aumenta el número de variables, se obtienen resultados en AUROC que son muy similares a los obtenidos con todas las variables. En cualquier caso, se debe estudiar si las diferencias son estadísticamente significativas entre los subconjuntos de (16,32) variables y el conjunto total.

Finalmente, como se puede ver en la Figura 15, el mejor modelo en tres de los cuatro conjuntos es SVM y, además, logra alcanzar valores superiores a 0.96 en AUROC.

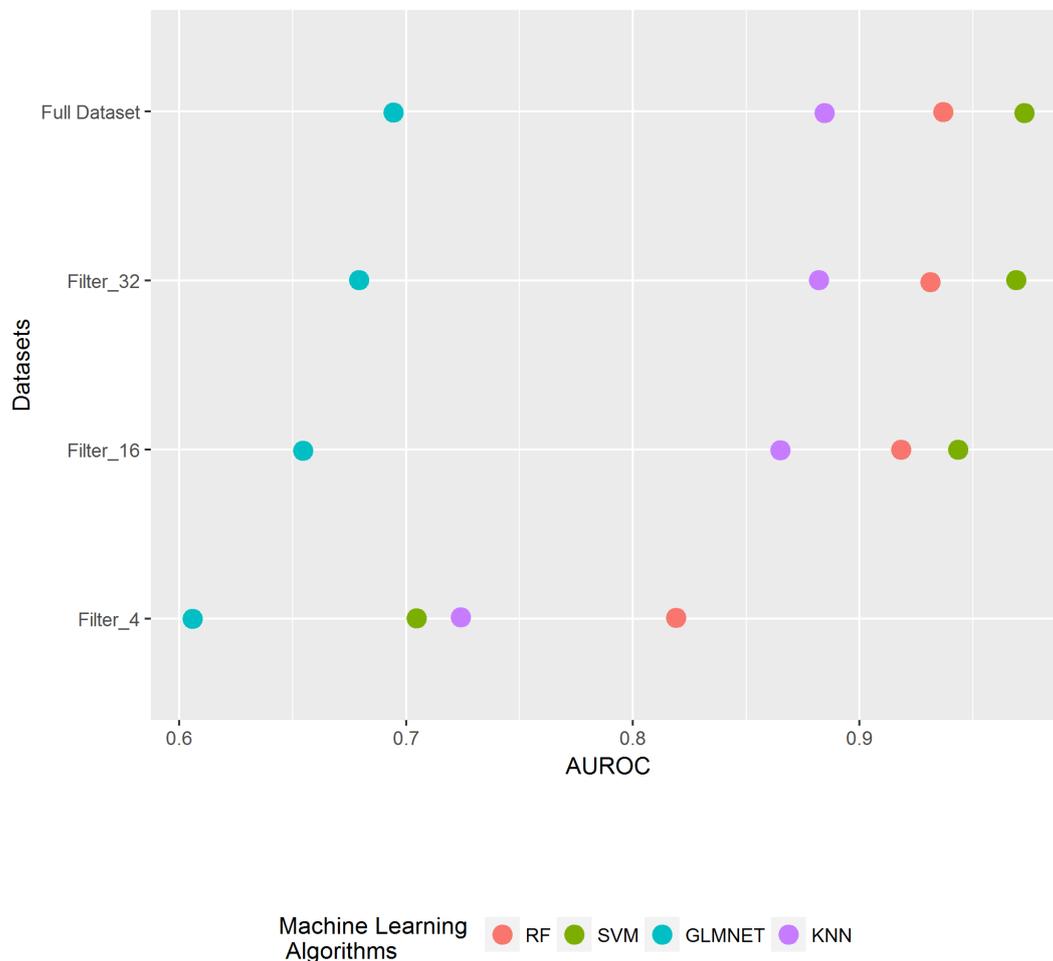


Figura 15. Resumen del desempeño. AUROC clasificación agregada de validación cruzada.

Sin embargo, tal como se mencionó anteriormente, esto no es suficiente y es necesario analizar el comportamiento de los modelos durante toda la fase experimental y verificar su estabilidad, como se puede ver en la Figura 16, la densidad de los resultados y su tendencia.

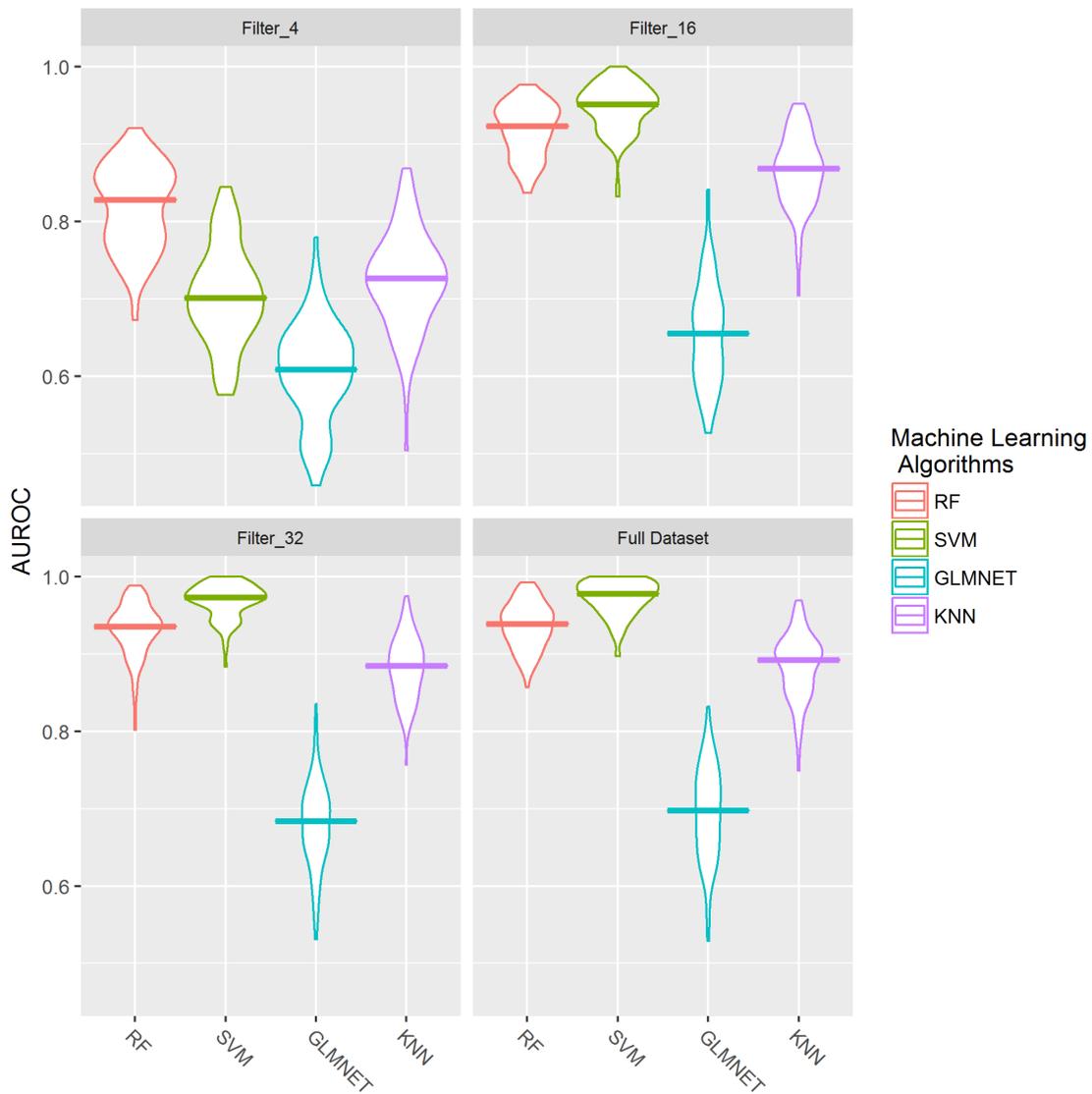


Figura 16. Estabilidad de los algoritmos de aprendizaje automático durante los experimentos.

Se puede observar en la Figura 16 que, si el número de variables es muy pequeño (4), los modelos están muy sesgados y existe una gran variabilidad porque realmente no hay suficiente información en los datos para encontrar un buen modelo de clasificación.

También es importante tener en cuenta que los resultados que se obtienen a medida que aumentan las variables, con tamaños de (16) y (32), esta variabilidad se reduce visualmente de manera significativa hasta alcanzar valores de promedio y desviación estándar muy similares a los obtenidos usando todas las variables.

Como se ha observado en las dos figuras anteriores, los mejores resultados en AUROC se han obtenido utilizando SVM. Ahora es el momento de verificar si la diferencia entre los tres modelos ganadores (SVM con 16, 32 y todas las variables) es significativa. Siguiendo la metodología experimental propuesta, se puede verificar si los resultados obtenidos por los tres mejores modelos siguen una distribución normal o no utilizando una prueba de Shapiro-Wilk, con un nivel de confianza  $\alpha = 0.05$ . Con los resultados en el trabajo de campo de la presente Tesis, se rechaza la hipótesis nula de que los resultados siguen una distribución con un valor  $W = 0.92328$  y un valor  $p - \text{value} = 2.722 \times 10^{-11}$ , por lo que se puede considerar que los datos no siguen esta distribución normal, como puede verse en la figura 17.

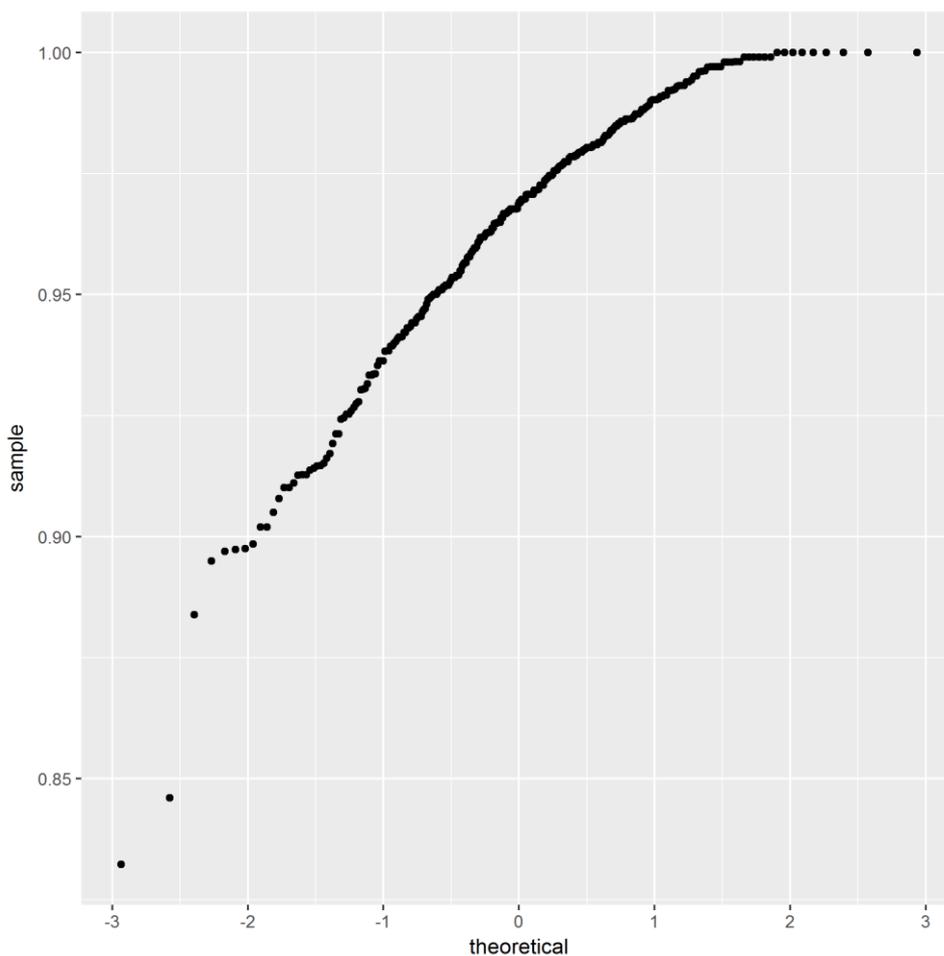


Figura 17. Gráfica Q-Q de valores observados *versus* valores esperados.

La Figura 18 permite visualizar la verificación de que, la falta de normalidad de los resultados se produce en los resultados obtenidos con SVM para los tres conjuntos de datos, no se debe a un subconjunto de los modelos generados.

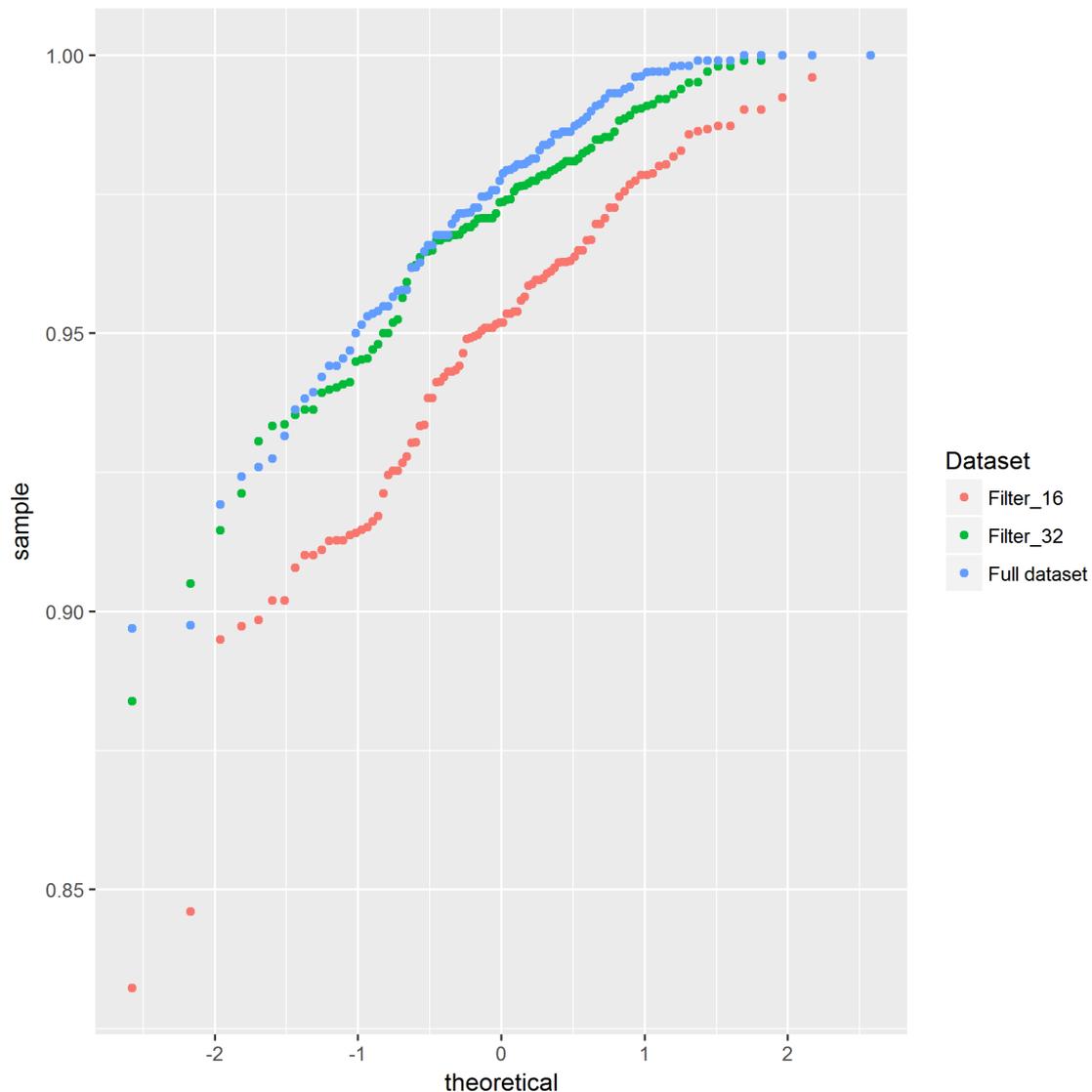


Figura 18. Gráfica Q-Q de los valores observados *versus* los esperados agrupados por conjunto de datos

Ahora se procede a realizar una prueba de Bartlett con un nivel de confianza y con la hipótesis nula para demostrar que los datos obtenidos presentan heteroscedasticidad. El resultado de la prueba indica que la hipótesis nula debe rechazarse con un valor de K-squared 11.64 de Barlett con 2 grados de libertad y p-valor = 0.002968. El resultado de

ambas pruebas indica que se debe realizar una prueba de Friedman no paramétrica, calculando el valor “p” con las pruebas de Iman y Davenport y con un nivel de confianza  $\alpha = 0.05$  asumiendo la hipótesis nula de que los resultados obtenidos en el trabajo de campo de la presente Tesis no tienen diferencias significativas entre ellos.

El resultado de esta prueba indica que la hipótesis no debe rechazarse con p – valor =  $6.78 \times 10^{-8}$ . La tabla 39 se presenta con el promedio de las clasificaciones de los tres modelos con las estadísticas Iman y Davenport (distribución F-con 2 y 198 grados de libertad: 19.57 y un p – valor =  $1.74 \times 10^{-8}$ ).

Cuanto menor sea el valor obtenido, mejor se considera el resultado. El modelo de control es el SVM que utiliza todas las variables.

Tabla 36. *Clasificación promedio obtenida con la prueba de Friedman y la extensión de Iman y Davenport.*

ML technique	Ranking
Full dataset	1.69
Filter_32	1.85
Filter_16	2.45

A través de una prueba de comparación múltiple, es necesario realizar un proceso de ajuste de los valores p obtenidos. En el presente caso, Finner realizará un ajuste que, según los autores, ofrece mejores resultados que los ajustes de Holm, Hochberg, Hommel, Holland y Rom. En este caso, el procedimiento de Finner rechaza las hipótesis que tienen un p-valor  $\leq 0.05$  y esto se verifica tanto para Filter\_16 (p-valor ajustado = 0.025) como Filter\_32 (p-valor ajustado = 0.05).

Estas pruebas sirven para verificar que el modelo ganador sea SVM y que haya utilizado todas las variables del estudio (conjunto de datos completo), por lo que se analizará la importancia de las variables en cada experimento realizado. Para hacer esto, como se mencionó anteriormente, se analizarán los valores de “w” de las diferentes variables y se calculará el promedio de la importancia de cada variable durante todos los experimentos. El resultado se puede ver en la Figura 19, dónde se observa que el metabolismo calórico basal (BMR), el riesgo de malnutrición según el MNA, el peso y la masa ósea, son las características más importantes resaltadas por el modelo ganador, SVM.

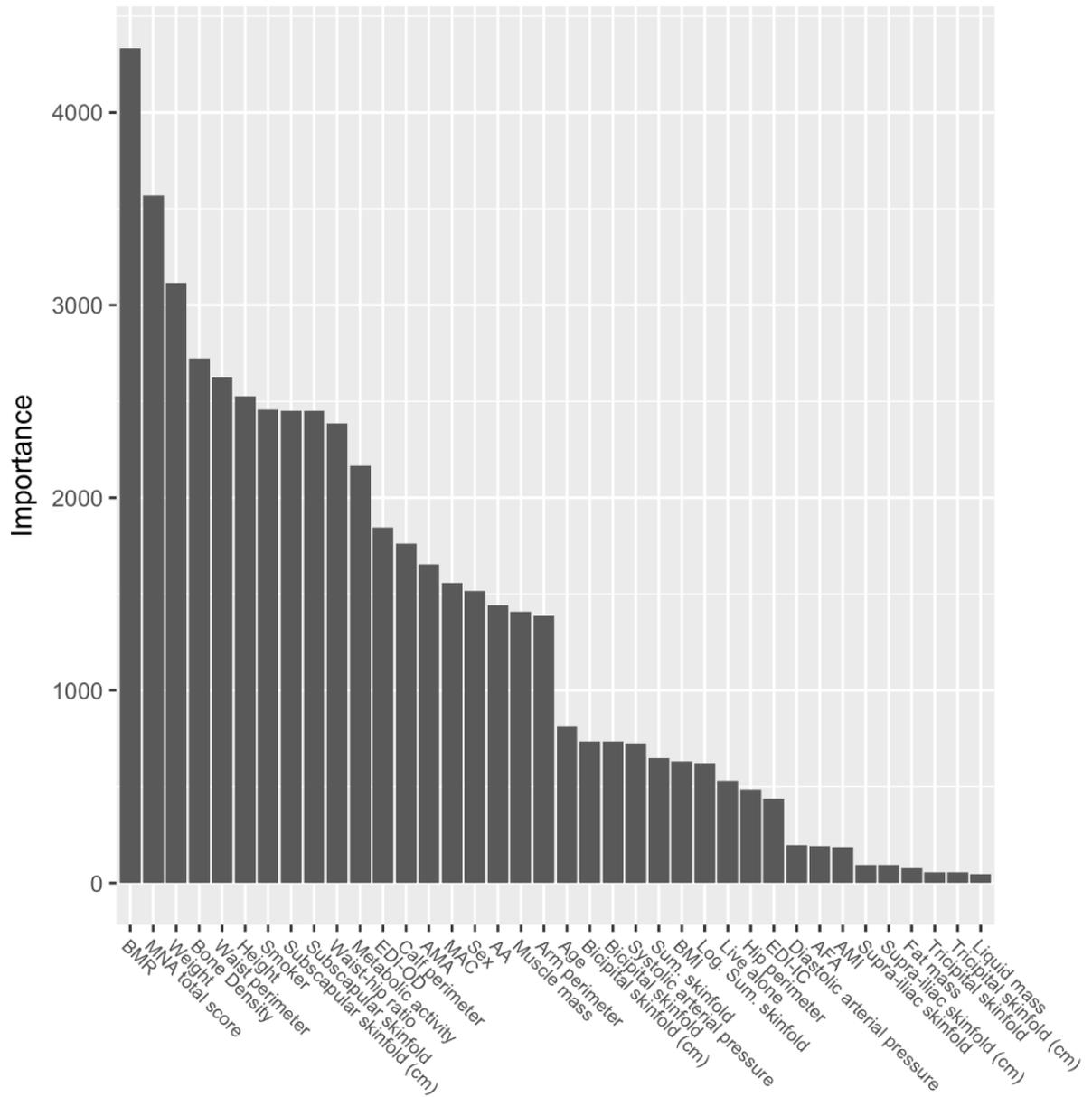


Figura 19. Importancia de la característica según el parámetro  $w$  del modelo ganador.



## 5. DISCUSIÓN Y CONCLUSIONES



Analizando los resultados obtenidos, se observan las diferencias y similitudes entre las técnicas estadísticas clásicas y las técnicas de aprendizaje automático en el análisis del estado nutricional y en la determinación de la adherencia a la dieta mediterránea en una muestra aleatoria poblacional de 503 habitantes del municipio de Cambre (A Coruña).

En primer lugar, se observa que la estadística clásica objetiva que las variables asociadas de manera independiente a la malnutrición son: vivir sólo (OR=2,92), mayor edad (OR=1,04), y perímetro de cintura menor (OR=0,96). Estando la adherencia a la dieta mediterránea en el límite de la significación estadística (OR=0,53;  $p=0,082$ ).

También se observa que, con las técnicas empleadas de ML, debido a la gran variabilidad de los resultados, se puede llegar a la conclusión de que ninguna de las variables da suficiente estabilidad como para que puedan tener significación a nivel clínico, con una insuficiente capacidad discriminadora que avale las variables asociadas a la malnutrición.

Estudios previos (159–161) han encontrado relevantes las mismas variables halladas con las técnicas clásicas, por lo que se han contrastado los resultados llegando a conclusiones similares. Además, la estadística clásica determina que las variables asociadas de manera independiente a una baja adherencia a la dieta mediterránea son: ser fumador (OR=2,0), mayor riesgo de malnutrición según el cuestionario MNA (OR=0,8), menor obsesión por la delgadez (OR=0,9), valores más elevados del pliegue subescapular (OR=1,78) y masa muscular (OR=1,2), valores más bajos de masa ósea (OR=0,3) y mayor metabolismo basal calórico (OR=1,0).

Con las técnicas de aprendizaje automático, que dan por ganador a la técnica SVM por ser la que ofrece unos resultados más significativos, permite concluir que las variables más relacionadas con el grado de adherencia a la dieta mediterránea son: el metabolismo basal calórico, el riesgo de malnutrición según el MNA, el peso y altura, la masa ósea, el índice cintura-cadera y el ser fumador.

Ambas aproximaciones, estadística clásica y ML, coinciden en establecer como variables más significativas: ser fumador, riesgo de malnutrición según la puntuación del MNA, valores de masa ósea y metabolismo basal calórico.

De igual manera que se ha realizado con el estado nutricional, para verificar si los resultados son relevantes y están de acuerdo con lo publicado anteriormente, se revisaron los artículos científicos más actuales publicados sobre el tema, en un intento por identificar el grado de adherencia a las variables más relacionadas. Los resultados de la búsqueda condujeron a estudios previos que también encontraron las variables identificadas como las más importantes, en particular el metabolismo basal calórico (162–165), el riesgo de malnutrición según la puntuación del MNA (166–168) la masa ósea (169–172) y si el paciente es fumador (167,173,174).

Por lo tanto, los resultados obtenidos en el trabajo de campo de la presente Tesis fueron debidamente contrastados en sus resultados, asegurando la capacidad de ambas técnicas de identificar patrones subyacentes en los datos. De acuerdo con el proceso de selección de características, los predictores restantes no se muestran relevantes.

## 6. FUTUROS DESARROLLOS



Actualmente, se están generando grandes cantidades de datos “ómicos”, con los que sería interesante probar los modelos propuestos en esta Tesis Doctoral para identificar si existe alguna firma genómica, proteómica, lipidómica, exposómica, ... que predisponga a los pacientes hacia un mejor estado nutricional y una alta adherencia a la dieta mediterránea.

Sería necesario recopilar más datos construyendo una especie de Big Data específico y probando sobre él los nuevos modelos de Deep Learning. Así como validar los resultados aquí obtenidos en otras cohortes. También resultaría de gran utilidad científica y asistencial generar un modelo web para validar los datos obtenidos y para comprobar nuevas propuestas.

Otro trabajo futuro que resultaría de interés siendo complementario al aquí presentado, sería desarrollar un sistema inteligente que aconseje y adopte la dieta mediterránea a las circunstancias, propias y de entorno, de cada individuo.

También sería muy interesante incluir entre los datos para analizar de cada individuo sus datos microbióticos, sobre todo los referidos a la microbiota bucofaríngea y de todo el tracto digestivo.

Cada día, en los últimos tiempos, se están encontrando y publicando evidencias de la importancia de los datos “ómicos” y de la microbiota en relación con el estado de salud de las personas, por ello se hace imprescindible tenerlos en cuenta para próximos estudios.

El problema hasta la fecha era no solo obtener la ingente cantidad de datos que ello supone, sino la propia complejidad de su análisis conjunto.

Las técnicas de Inteligencia Artificial como las utilizadas en la presente Tesis Doctoral y otras como el “Deep Learning”, la Programación genética, etc. hacen posible y facilitan estos complejos análisis de grandes cantidades de datos complejos, permitiendo extraer nuevo conocimiento de los mismos y haciendo plausible una aplicación clínica de los resultados que ayudan a mejorar la calidad de vida de las personas y la sostenibilidad de los sistemas públicos de salud.



## REFERENCIAS

1. García De Lorenzo A, Álvarez J, Calvo M V., De Ulíbarri JI, Del Río J, Galbán C, et al. Conclusiones del II Foro de Debate SENPE sobre desnutrición hospitalaria. *Nutr Hosp.* 2005;20(2):82–7.
2. Organización Mundial de la Salud O, Ministerio de Sanidad Servicios Sociales e Igualdad G de E. CIE 9 MC. Clasificación Internacional de Enfermedades [Internet]. 2014. 1149–1150 p. Available from: [https://eciemaps.mspsi.es/ecieMaps/browser/index\\_9\\_mc.html](https://eciemaps.mspsi.es/ecieMaps/browser/index_9_mc.html)
3. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. *Jama.* 2004;291(10):1238–45.
4. Banegas JR, López-García E, Gutiérrez-Fisac JL, Guallar-Castillón P, Rodríguez-Artalejo F. A simple estimate of mortality attributable to excess weight in the European Union. *Eur J Clin Nutr* [Internet]. 2003;57(2):201–8. Available from: <https://doi.org/10.1038/sj.ejcn.1601538>
5. OMS | Informe sobre la situación mundial de las enfermedades no transmisibles 2014. WHO [Internet]. 2016 [cited 2019 Jul 22]; Available from: [https://www.who.int/nmh/publications/ncd-status-report-2014/es/#.XTWBHfb\\_oMg.mendeley](https://www.who.int/nmh/publications/ncd-status-report-2014/es/#.XTWBHfb_oMg.mendeley)
6. OMS | 10 datos sobre la obesidad. WHO [Internet]. 2017 [cited 2019 Jul 21]; Available from: <https://www.who.int/features/factfiles/obesity/es/#.XTSX1JSr0Jo.mendeley>
7. Malo-serrano M, M NC, D DP, Malo-serrano M. La obesidad en el mundo Obesity in the world. 2017;173–8.
8. Europea E, Nacional E, Nacional LE, Europea E, Nacional LE, En S. Salud 4 . 6 Determinantes de salud ( sobrepeso , consumo de fruta y verdura , tipo de lactancia , actividad física ) Salud. 2018;(Imc):1–10.
9. Aranceta-Bartrina J, Pérez-Rodrigo C, Alberdi-Aresti G, Ramos-Carrera N, Lázaro-Masedo S. Prevalencia de obesidad general y obesidad abdominal en la población adulta española (25–64 años) 2014–2015: estudio ENPE. *Rev Española Cardiol.* 2016;69(6):579–87.
10. BMI OC. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults. *NOE Initiat.* 1998;6(Suppl 2):51S-209S.
11. Moyer VA. Screening for and management of obesity in adults: US Preventive Services Task Force recommendation statement. *Ann Intern Med.* 2012;157(5):373–8.
12. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med.* 2006;3(11):e442.

13. Wilson PWF, D'Agostino RB, Sullivan L, Parise H, Kannel WB. Overweight and obesity as determinants of cardiovascular risk: the Framingham experience. *Arch Intern Med.* 2002;162(16):1867–72.
14. Sanidad MDE, Igualdad SSE. Plan de calidad para el Sistema Nacional de Salud 2010. Plan Calid para el Sist Nac Salud [Internet]. 2010; Available from: <http://www.msssi.gob.es/organizacion/sns/planCalidadSNS/pncalidad.htm%5Cn> <http://www.msssi.gob.es/organizacion/sns/planCalidadSNS/pdf/pncalidad/PlanCalidad2010.pdf>
15. Luengo Fernández E, Ordóñez Rubio B, Bergua Martínez C, Laclaustra Gimeno M. Obesidad, dislipemia y síndrome metabólico. *Rev Española Cardiol Supl* [Internet]. 2005 Jan 1 [cited 2019 Oct 21];5(4):21D-29D. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S1131358705741170>
16. Detsky AS, Baker JP, Johnston N, Whittaker S, Mendelson RA, Jeejeebhoy KN. What is subjective global assessment of nutritional status? *J Parenter Enter Nutr.* 1987;11(1):8–13.
17. Kondrup J, Allison SP, Elia M, Vellas B, Plauth M. ESPEN guidelines for nutrition screening 2002. *Clin Nutr.* 2003;22(4):415–21.
18. Kaiser MJ, Bauer JM, Ramsch C, Uter W, Guigoz Y, Cederholm T, et al. Validation of The Mini Nutritional Assessment - Short Form (MNA-SF) : A Practical Tool for Identification of Nutritional Status. *J Nutr Health Aging.* 2009;13(January):782–8.
19. De Ulíbarri JI, González-Madroño A, de Villar NGP, González P, González B, Mancha A, et al. CONUT: a tool for controlling nutritional status. First validation in a hospital population. *Nutr Hosp.* 2005;20(1):38–45.
20. De la Montaña Miguélez J, Areal Salve C, Míguez Bernárdez M. Evaluación del riesgo nutricional mediante el MNA en una población anciana no institucionalizada. *Arch Latinoam Nutr.* 2009;59(4):390–5.
21. Cuervo M, García A, Ansorena D, Sánchez-Villegas A, Martínez-González M, Astiasarán I, et al. Nutritional assessment interpretation on 22 007 Spanish community-dwelling elders through the Mini Nutritional Assessment test. *Public Health Nutr* [Internet]. 2009 Jan 16 [cited 2017 Nov 2];12(01):82. Available from: [http://www.journals.cambridge.org/abstract\\_S136898000800195X](http://www.journals.cambridge.org/abstract_S136898000800195X)
22. Gil-Montoya JA, Subirá C, Ramón JM, González-Moles MA. Oral health-related quality of life and nutritional status. *J Public Health Dent.* 2008;68(2):88–93.
23. Jiménez Sanz MJ, Sola Villafranca J, Pérez Ruiz C, Turienzo Llata MJ, Larrañaga Lavin G, Mancebo Santamaría MA, et al. Estudio del estado nutricional de los ancianos de Cantabria. *Nutr Hosp.* 2011;26(2):345–54.
24. Míguez Bernárdez M, De la Montaña Miguélez J, González Carnero J, González Rodríguez M. Concordancia entre la autopercepción de la imagen corporal y el estado nutricional en universitarios de Orense. *Nutr Hosp.* 2011;26(3):472–9.

25. Mujica VE, Leiva EM, Rojas EM, Díaz N, Icaza G, Palomo IG. Discordancia en autopercepción de peso en población adulta de Talca. *Rev Med Chil*. 2009;137(1):76–82.
26. Rosen JC. Tratamiento cognitivo-conductual para el trastorno dismórfico corporal. *Man para el Trat Cogn los Trastor psicológicos*. 1997;1(1):30–58.
27. Garner DM, Olmstead MP, Polivy J. Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *Int J Eat Disord*. 1983;2(2):15–34.
28. Phillips KA. *Body dysmorphic disorder modification of the YBOCS, McLean version*. Belmont Massachusetts. 1993;
29. Cooper PJ, Taylor MJ, Cooper Z, Fairbum CG. The development and validation of the Body Shape Questionnaire. *Int J Eat Disord*. 1987;6(4):485–94.
30. Perugi G, Giannotti D, Di Vaio S, Frare F, Sacttoni M, Cassano GB. Fluvoxamine in the treatment of body dysmorphic disorder (dysmorphophobia). *Int Clin Psychopharmacol*. 1996;11(4):247–54.
31. Cash TF. The great American shape-up: Body-image survey report. *Psychol Today*. 1986;20:30–7.
32. Rucker III CE, Cash TF. Body images, body-size perceptions, and eating behaviors among African-American and white college women. *Int J Eat Disord*. 1992;12(3):291–9.
33. Cash TF, Labarge AS. Development of the Appearance Schemas Inventory: A new cognitive body-image assessment. *Cognit Ther Res*. 1996;20(1):37–50.
34. Rosen JC, Srebnik D, Saltzberg E, Wendt S. Development of a body image avoidance questionnaire. *Psychol Assess A J Consult Clin Psychol*. 1991;3(1):32.
35. Ríos PB. Patricia Bolaños Ríos. *Trastornos de la Conducta Alimentaria* 9 (2009) 956-972. 2009;9:956–72.
36. Dussailant C, Echeverría G, Urquiaga Inés, Velasco Nicolás, Rigotti Attilio. Evidencia actual sobre los beneficios de la dieta mediterránea en salud. artículo revisión *rev Med chile* *Rev Med Chile* [Internet]. 2016;144(144):1044–52. Available from: <http://www.scielo.cl/pdf/rmc/v144n8/art12.pdf>
37. Keys A. Coronary heart disease in seven countries. *Circulation*. 1970;41(1):186–95.
38. Keys A. Mediterranean diet and public health: personal reflections. *Am J Clin Nutr*. 1995;61(6):1321S-1323S.
39. Simopoulos AP. The Mediterranean diets: what is so special about the diet of Greece? The scientific evidence. *J Nutr*. 2001;131(11):3065S-3073S.
40. Sesso HD, Liu S, Gaziano JM, Buring JE. Dietary lycopene, tomato-based food products and cardiovascular disease in women. *J Nutr*. 2003;133(7):2336–41.

41. Tresserra-Rimbau A, Rimm EB, Medina-Remón A, Martínez-González MA, De la Torre R, Corella D, et al. Inverse association between habitual polyphenol intake and incidence of cardiovascular events in the PREDIMED study. *Nutr Metab Cardiovasc Dis.* 2014;24(6):639–47.
42. Threapleton DE, Greenwood DC, Evans CEL, Cleghorn CL, Nykjaer C, Woodhead C, et al. Dietary fibre intake and risk of cardiovascular disease: systematic review and meta-analysis. *Bmj.* 2013;347:f6879.
43. Wu Y, Qian Y, Pan Y, Li P, Yang J, Ye X, et al. Association between dietary fiber intake and risk of coronary heart disease: A meta-analysis. *Clin Nutr.* 2015;34(4):603–11.
44. Slavin JL. Dietary fiber and body weight. *Nutrition.* 2005;21(3):411–8.
45. Escurriol V, Cofán M, Serra M, Bulló M, Basora J, Salas-Salvadó J, et al. Serum sterol responses to increasing plant sterol intake from natural foods in the Mediterranean diet. *Eur J Nutr.* 2009;48(6):373–82.
46. Covas M-I, de la Torre R, Fitó M. Virgin olive oil: a key food for cardiovascular risk protection. *Br J Nutr.* 2015;113(S2):S19–28.
47. Simopoulos AP. The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. *Exp Biol Med.* 2008;233(6):674–88.
48. Fontana L, Bermudez-Brito M, Plaza-Diaz J, Munoz-Quezada S, Gil A. Sources, isolation, characterisation and evaluation of probiotics. *Br J Nutr.* 2013;109(S2):S35–50.
49. Bach-Faig A, Fuentes-Bol C, Ramos D, Carrasco JL, Roman B, Bertomeu IF, et al. The Mediterranean diet in Spain: adherence trends during the past two decades using the Mediterranean Adequacy Index. *Public Health Nutr.* 2011;14(4):622–8.
50. Knoops KTB, Fidanza F, Alberti-Fidanza A, Kromhout D, Van Staveren WA. Comparison of three different dietary scores in relation to 10-year mortality in elderly European subjects: the HALE project. *Eur J Clin Nutr.* 2006;60(6):746.
51. Moreiras GV, Torres JMÁ, Vives CC, de la Calle S del P, Moreno ER. Valoración de la dieta española de acuerdo al Panel de Consumo Alimentario. *Distrib y Consum.* 2009;19(105):20–8.
52. Trichopoulou A, Costacou T, Bamia C, Trichopoulos D. Adherence to a Mediterranean diet and survival in a Greek population. *N Engl J Med.* 2003;348(26):2599–608.
53. Martínez-González MA, García-López M, Bes-Rastrollo M, Toledo E, Martínez-Lapiscina EH, Delgado-Rodríguez M, et al. Mediterranean diet and the incidence of cardiovascular disease: A Spanish cohort. *Nutr Metab Cardiovasc Dis.* 2011;21(4):237–44.

54. Martínez-González MÁ, De la Fuente-Arrillaga C, Nunez-Cordoba JM, Basterra-Gortari FJ, Beunza JJ, Vazquez Z, et al. Adherence to Mediterranean diet and risk of developing diabetes: prospective cohort study. *Bmj*. 2008;336(7657):1348–51.
55. Bonaccio M, Di Castelnuovo A, Costanzo S, Persichillo M, De Curtis A, Donati MB, et al. Adherence to the traditional Mediterranean diet and mortality in subjects with diabetes. Prospective results from the MOLI-SANI study. *Eur J Prev Cardiol*. 2016;23(4):400–7.
56. Trichopoulou A, Bamia C, Trichopoulos D. Anatomy of health effects of Mediterranean diet: Greek EPIC prospective cohort study. *Bmj* [Internet]. 2009;338(jun23 2):b2337–b2337. Available from: <http://www.bmj.com/cgi/doi/10.1136/bmj.b2337>
57. Sofi F, Cesari F, Abbate R, Gensini GF, Casini A. Adherence to Mediterranean diet and health status: meta-analysis. *Bmj*. 2008;337:a1344.
58. Giacosa A, Barale R, Bavaresco L, Gatenby P, Gerbi V, Janssens J, et al. Cancer prevention in Europe: the Mediterranean diet as a protective choice. *Eur J Cancer Prev*. 2013;22(1):90–5.
59. Sofi F, Macchi C, Casini A. Mediterranean diet and minimizing neurodegeneration. *Curr Nutr Rep*. 2013;2(2):75–80.
60. Shapiro SC. *Encyclopedia of artificial intelligence second edition*. John; 1992.
61. El Naqa I, Murphy MJ. What is machine learning? In: *Machine Learning in Radiation Oncology*. Springer; 2015. p. 3–11.
62. Warren M, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys*. 1943;5(4):115–33.
63. Warner HR, Toronto AF, Veasey LG, Stephenson R. A Mathematical Approach to Medical Diagnosis: Application to Congenital Heart Disease. *JAMA* [Internet]. 1961 Jul 22;177(3):177–83. Available from: <https://doi.org/10.1001/jama.1961.03040290005002>
64. Yu VL, Fagan LM, Wraith SM, Clancey WJ, Scott AC, Hannigan J, et al. Antimicrobial selection by a computer. A blinded evaluation by infectious diseases experts. *Jama*. 1979;242(12):1279–82.
65. Siegel JD, Parrino TA. Computerized diagnosis: implications for clinical education. *Med Educ*. 1988;22(1):47–54.
66. Fernández SP, Díaz SP, Calviño BL, Santamaría PG, Pillado TS, Monreal FA, et al. Diagnosis delay and follow-up strategies in colorectal cancer. Prognosis implications: a study protocol. *BMC Cancer*. 2010;10(1):528.
67. Pereira M, Curra A, Rivas R, Pereira J, Banos G, Teueiro J, et al. Computer aided monitoring system of intensive care unit patients. *WSEAS Trans Inf Sci Appl*. 2007;4(1):78–84.

68. Rivero D, Dorado J, Rabuñal J, Pazos A. Using genetic programming for artificial neural network development and simplification. In: Proceedings of the 5th WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics. Citeseer; 2006. p. 65–71.
69. Berca MN, Duardo-Sanchez A, González-Díaz H, Pazos A, Munteanu CR. 10. Markov entropy for biology, parasitology, linguistic, technology, social and law networks. 2011;
70. Audi R, Audi P. The Cambridge dictionary of philosophy. Vol. 584. Cambridge university press Cambridge; 1999.
71. Chen H-L, Yang B, Wang G, Wang S-J, Liu J, Liu D-Y. Support vector machine based diagnostic system for breast cancer using swarm intelligence. *J Med Syst.* 2012;36(4):2505–19.
72. Tenório JM, Hummel AD, Cohrs FM, Sdepanian VL, Pisa IT, de Fátima Marin H. Artificial intelligence techniques applied to the development of a decision–support system for diagnosing celiac disease. *Int J Med Inform.* 2011;80(11):793–802.
73. González R, Barrientos A, Toapanta M, Del Cerro J. Aplicación de las Máquinas de Soporte Vectorial (SVM) al diagnóstico clínico de la Enfermedad de Párkinson y el Temblor Esencial. *RIAI - Rev Iberoam Autom e Inform Ind [Internet].* 2017;14(4):394–405. Available from: <http://dx.doi.org/10.1016/j.riai.2017.07.005>
74. Krawczyk B, Simić D, Simić S, Woźniak M. Automatic diagnosis of primary headaches by machine learning methods. *Open Med.* 2013;8(2):157–65.
75. Weng SF, Vaz L, Qureshi N, Kai J. Prediction of premature all-cause mortality: A prospective general population cohort study comparing machine-learning and standard epidemiological approaches. Rezaei-Hachesu P, editor. *PLoS One [Internet].* 2019 Mar 27 [cited 2019 Aug 30];14(3):e0214365. Available from: <http://dx.plos.org/10.1371/journal.pone.0214365>
76. Watson I. Case-based reasoning is a methodology not a technology. In: *Research and Development in Expert Systems XV.* Springer; 1999. p. 213–23.
77. Samarghitean C, Ortutay C, Vihinen M. Systematic classification of primary immunodeficiencies based on clinical, pathological, and laboratory parameters. *J Immunol.* 2009;183(11):7569–75.
78. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Ser B Stat Methodol.* 2005;67:301–20.
79. Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics.* 1970;12(1):55–67.
80. Tibshirani R. Regression Selection and Shrinkage via the Lasso. Vol. 58, *Journal of the Royal Statistical Society B.* 1994. p. 267–88.

81. Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature*. 2012;483(7391):570.
82. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14(10):3156.
83. Whelan R, Watts R, Orr CA, Althoff RR, Artiges E, Banaschewski T, et al. Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature*. 2014;512(7513):185.
84. Charlson ME, Pompei P, Ales KL, MacKenzie R. A new method of classifying prognostic in longitudinal studies: development and validation. Vol. 40, *Journal of Chronic Diseases*. 1987. p. 373–83.
85. Charlson M, Szatrowski TP, Peterson J, Gold J. Validation of a combined comorbidity index. *J Clin Epidemiol*. 1994;47(11):1245–51.
86. Charlson ME, Charlson RE, Peterson JC, Marinopoulos SS, Briggs WM, Hollenberg JP. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. *J Clin Epidemiol [Internet]*. 2008 Dec 1 [cited 2018 Apr 20];61(12):1234–40. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18619805>
87. Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom K .O. The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br J Addict*. 1991;86(9):1119–27.
88. Becoña Iglesias E, Lorenzo Pontevedra MC. Evaluación de la conducta de fumar. *Adicciones*. 2004;16(SUPPL. 2):201–26.
89. Carrasco T, Luna M, Vila J. Validez del “Fageström Tolerance Questionnaire” como medida de dependencia física de la nicotina: una revisión. Vol. 19, *Revista Española de Drogodependencias*. 1994. p. 3–14.
90. Córdoba R, Martín C, Casas R, Barberá C, Botaya M, Hernández A, et al. Valor de los cuestionarios breves en la predicción del abandono del tabaco en atención primaria. *Atención Primaria [Internet]*. 2000 [cited 2018 Apr 22];25(1):32–6. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0212656700784605>
91. Lee EW, D’Alonzo GE. Cigarette smoking, nicotine addiction, and its pharmacologic treatment. *Arch Intern Med*. 1993;153(1):34–48.
92. Alastrué A, Rull M, Camps I SJ. Nuevas normas y consejos en la valoración de los parámetros antropométricos en nuestra población: índice adiposo muscular, índices ponderales y tablas de percentiles de los datos antropométricos útiles en una valoración nutricional. *Med Clin*. 1988;(91):223–36.
93. Alastrué A, Sitges A, Jaurrieta E SA. Valoración de los parámetros antropométricos en nuestra población. *Med Clin*. 1982;(10):407–15.

94. Bjorntorp P, Bray GA, Carroll KK, Chuchalin A, Dietz WH, Ehrlich GE, et al. Obesity: Preventing and Managing the Global Epidemic. WHO Tech Rep Ser. 2000;253.
95. Moreno González MI. Circunferencia de cintura: una medición importante y útil del riesgo cardiometabólico. *Rev Chil Cardiol*. 2010;29(1):85–7.
96. Martínez Usó I, Civera Andrés M. Protocolo diagnóstico de la malnutrición. *Med - Programa Form Médica Contin Acreditado* [Internet]. 2002 Jan 1 [cited 2018 Apr 22];8(87):4717–9. Available from: <https://www.sciencedirect.com/science/article/pii/S0304541202708743>
97. Rolland Y, Lauwers-Cances V CM y cols. Sarcopenia, calf circumference, and physical function of elderly women: a cross-sectional study. *J Am Geriatr Soc*. 2003;51(8):1120–4.
98. Barbosa Murillo JAP, Rodríguez M NG, Hernández H De Valera YM, Hernández H RA, Herrera M HA. Masa muscular, fuerza muscular y otros componentes de funcionalidad en adultos mayores institucionalizados de la Gran Caracas-Venezuela. *Nutr Hosp*. 2007;22(5):578–83.
99. Bonnefoy M, Jauffret M, Kostka T JJ. Usefulness of calf circumference measurement in assessing the nutritional state of hospitalized elderly people. *Gerontology*. 2002;48(3):162–9.
100. Norman K, Smoliner C, Kilbert A, Valentini L, Lochs H, Pirlich M. Disease-related malnutrition but not underweight by BMI is reflected by disturbed electric tissue properties in the bioelectrical impedance vector analysis. *Br J Nutr* [Internet]. 2008 Sep [cited 2017 Nov 2];100(03):590–5. Available from: [http://www.journals.cambridge.org/abstract\\_S0007114508911545](http://www.journals.cambridge.org/abstract_S0007114508911545)
101. Norman K, Smoliner C, Valentini L, Lochs H, Pirlich M. Is bioelectrical impedance vector analysis of value in the elderly with malnutrition and impaired functionality? *Nutrition*. 2007;23(7–8):564–9.
102. Wirth R, Miklis P. Die Bioelektrische Impedanzanalyse in der Diagnostik der Malnutrition. *Z Gerontol Geriatr* [Internet]. 2005;38(5):315–21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16244815>
103. Gmbh B. Diagnose-Waage Gebrauchsanweisung.....2. 49(0).
104. GUIGOZ, Y. Mini Nutritional Assessment : A practical assessment tool for grading the nutritional state of elderly patients. *Facts Res Geyontol* [Internet]. 1994 [cited 2018 Apr 24];4(2):15–59. Available from: <http://ci.nii.ac.jp/naid/10030728344/en/>
105. Ruiz-López MD, Artacho R, Oliva P, Moreno-Torres R, Bolaños J, de Teresa C, et al. Nutritional risk in institutionalized older women determined by the Mini Nutritional Assessment test: what are the main factors? *Nutrition* [Internet]. 2003 Sep 1 [cited 2018 Apr 24];19(9):767–71. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12921887>

106. Salvà A. Un nuevo instrumento para la valoración nutricional en geriatría el "Mini Nutritional Assessment"(MNA). *Rev Multidiscip Gerontol.* 1996;6(4):319–28.
107. Guigoz Y, Vellas B, Garry PJ. Assessing the nutritional status of the elderly: The Mini Nutritional Assessment as part of the geriatric evaluation. *Nutr Rev.* 1996;54(1):S59.
108. Helsing E. Traditional diets and disease patterns of the Mediterranean, circa 1960. *Am J Clin Nutr.* 1995;61(6):1329S-1337S.
109. Willett WC, Sacks F, Trichopoulou A, Drescher G, Ferro-Luzzi A, Helsing E, et al. Mediterranean diet pyramid: a cultural model for healthy eating. *Am J Clin Nutr.* 1995;61(6):1402S-1406S.
110. Trichopoulou A, Lagiou P. Healthy traditional Mediterranean diet: an expression of culture, history, and lifestyle. *Nutr Rev.* 1997;55(11):383–9.
111. Díaz I, Gascón E, Lázaro S, Maximiano C. Guía de la alimentación mediterránea. Ed Empres Pública Desarro Agrar y Pesq Cons Agric y Pesca Junta Andalucía. 2007;
112. Sanchez-Villegas A, Bes-Rastrollo M, Martinez-Gonzalez MA, Serra-Majem L. Adherence to a Mediterranean dietary pattern and weight gain in a follow-up study: the SUN cohort. *Int J Obes.* 2006;30(2):350–8.
113. Hu FB. The Mediterranean diet and mortality-olive oil and beyond. *N Engl J Med.* 2003;348(26):2595–6.
114. Tur JA, Serra-Majem L, Romaguera D, Pons A. Does the diet of the Balearic population, a Mediterranean type diet, still provide adequate antioxidant nutrient intakes? *Eur J Nutr.* 2005;44(4):204–13.
115. Kris-Etherton P, Eckel RH, Howard B V, Jeor SS, Bazzarre TL. Lyon diet heart study: benefits of a Mediterranean-Style, National Cholesterol Education Program/American Heart Association Step I dietary pattern on cardiovascular disease. *Circulation.* 2001;103(13):1823–5.
116. Ferrara LA, Raimondi AS, d'Episcopo L, Guida L, Russo A Dello, Marotta T. Olive oil and reduced need for antihypertensive medications. *Arch Intern Med.* 2000;160(6):837–42.
117. Estruch R, Martínez-González MA, Corella D, Salas-Salvadó J, Ruiz-Gutiérrez V, Covas MI, et al. Effects of a Mediterranean-style diet on cardiovascular risk factors: a randomized trial. *Ann Intern Med.* 2006;145(1):1–11.
118. Hu FB, Rimm EB, Stampfer MJ, Ascherio A, Spiegelman D, Willett WC. Prospective study of major dietary patterns and risk of coronary heart disease in men-. *Am J Clin Nutr.* 2000;72(4):912–21.
119. Terry P, Hu FB, Hansen H, Wolk A. Prospective study of major dietary patterns and colorectal cancer risk in women. *Am J Epidemiol.* 2001;154(12):1143–9.

120. Trichopoulou, Naska, Antoniou, Friel, Trygg, Turrini. Vegetable and fruit: the evidence in their favour and the public health perspective. *Int J Vitam Nutr Res.* 2003;73(2):63–9.
121. Trichopoulou A, Orfanos P, Norat T, Bueno-de-Mesquita B, Ocké MC, Peeters PHM, et al. Modified Mediterranean diet and survival: EPIC-elderly prospective cohort study. *BMJ [Internet].* 2005;330(7498):991. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15820966>
122. Marco JH, Perpiña C, Baños R, Botella C, Borra C, Gallego MJ, et al. La Importancia de la imagen corporal en los trastornos alimentarios. 2004;
123. Corral S, González M, Pereña J, Seisdedos N. Adaptación española del Inventario de trastornos de la conducta alimentaria. EDI-2 Inventar Trastor la Conduct Aliment Manual Madrid TEA. 1998;69–80.
124. Garner DM. EDI-2: Inventario de Trastornos de la Conducta Alimentaria. Manual. MADRID; 1998.
125. Bruch H. Perceptual and conceptual disturbances in anorexia nervosa. *Psychosom Med.* 1962;24(2):187–94.
126. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2016.
127. Bischl B, Lang M, Kotthoff L, Schiffner J, Richter J, Jones Z, et al. *mlr: Machine Learning in R.* 2016.
128. Bacon F. Método inductivo-deductivo . 2003;1–11.
129. García S, Luengo J, Herrera F. Tutorial on practical tips of the most influential data preprocessing algorithms in data mining. *Knowledge-Based Syst.* 2016;98:1–29.
130. Dasu T, Johnson T. *Exploratory data mining and data cleaning.* Vol. 479. John Wiley & Sons; 2003.
131. Mosteller F, Tukey JW. Data analysis, including statistics. *Handb Soc Psychol.* 1968;2:80–203.
132. Oommen T, Misra D, Twarakavi NKC, Prakash A, Sahoo B, Bandopadhyay S. An objective analysis of support vector machine based classification for remote sensing. *Math Geosci.* 2008;40(4):409–24.
133. James G, Witten D, Hastie T, Tibshirani R. *An introduction to statistical learning.* Vol. 112. Springer; 2013.
134. Chaparro J, Giraldo BG, Rondón S. Evaluación del clasificador Naïve Bayes como herramienta de diagnóstico en Unidades de Cuidado Intensivo. *Rev Tecnol.* 2013;12(2):87–93.
135. Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics.* 2007 Oct;23(19):2507–17.

136. Kohavi R, John GH. Wrappers for feature subset selection. *Artif Intell.* 1997;97(1–2):273–324.
137. Guyon I, Elisseeff A. An introduction to variable and feature selection. *J Mach Learn Res.* 2003;3(Mar):1157–82.
138. Al B. Lum and P. L. Langley. Selection of relevant features and examples in machine learning. *Artif Intell (Special Issue Relev.* 1997;92(1–2):245–71.
139. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett.* 2006;27(8):861–74.
140. García S, Fernández A, Luengo J, Herrera F. Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. *Inf Sci (Ny).* 2010;180(10):2044–64.
141. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika.* 1965;52(3–4):591–611.
142. Bartlett MS. Properties of Sufficiency and Statistical Tests. *Proc R Soc Lond A Math Phys Sci.* 1937;160(901):268–82.
143. Friedman M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J Am Stat Assoc.* 1937;32(200):675–701.
144. Scholkopf B, Smola AJ, Schölkopf B. *Learning with kernels.* 2002;
145. Schölkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RC. Estimating the Support of a High-Dimensional Distribution. *Neural Comput [Internet].* 2001 Jul [cited 2018 Apr 17];13(7):1443–71. Available from: <http://www.mitpressjournals.org/doi/10.1162/089976601750264965>
146. Chih-Wei Hsu, Chih-Chung Chang and C-JL. A Practical Guide to Support Vector Classification. *BJU Int [Internet].* 2008;101(1):1396–400. Available from: <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>
147. Carmona Suárez EJ. Máquinas de Vectores Soporte (SVM). Dpto Intel Artif ETS Ing Inforática, Univ Nac Educ a Distancia [Internet]. 2014;1–25. Available from: [http://www.ia.uned.es/~ejcarmona/publicaciones/\[2013-Carmona\] SVM.pdf](http://www.ia.uned.es/~ejcarmona/publicaciones/[2013-Carmona] SVM.pdf)
148. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybern.* 1980;36(4):193–202.
149. Reyna A, Patronés R De. SVM de una clase : aplicación a detección de novedad. 2009;
150. Burges CJC. A Tutorial on Support Vector Machines for Pattern Recognition. Fayyad U, editor. *Data Min Knowl Discov.* 1998;2(2):121–167.
151. Campbell C, Ying Y. Learning with support vector machines. *Synth Lect Artif Intell Mach Learn.* 2011;5(1):1–95.

152. Shawe-Taylor J, Cristianini N. Kernel methods for pattern analysis. Cambridge university press; 2004.
153. Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines: And Other Kernel-based Learning Methods. New York, NY, USA: Cambridge University Press; 2000.
154. Mercer J. Xvi. functions of positive and negative type, and their connection the theory of integral equations. Philos Trans R Soc London Ser A, Contain Pap a Math or Phys character. 1909;209(441–458):415–46.
155. Breiman L. Random Forest. Mach Learn. 2001;45:5–32.
156. García C, Gómez I. Algoritmos de aprendizaje: knn & kmeans. Univ Carlos III Madrid [Internet]. 2006;1–8. Available from: <http://www.it.uc3m.es/jvillena/irc/practicas/08-09/06.pdf>
157. Cambroner CG, Moreno IG. Algoritmos de aprendizaje: knn & kmeans. Inteligencia en Redes Comun Univ Carlos III Madrid. 2006;23.
158. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. J R Stat Soc Ser B. 1996;58:267–88.
159. Montejano Lozoya a R, Ferrer Diego RM, Clemente Marin G, Martinez-Alzamora N, Sanjuan Quiles A, Ferrer Ferrandiz E. Factores asociados al riesgo nutricional en adultos mayores autónomos no institucionalizados. Nutr Hosp. 2014;30(4):858–69.
160. Fernández Rodríguez V, Fernández Barón A, Ordoño Ceba M. Salud sexual y envejecimiento cognitivo [Internet]. Salud, alimentación y sexualidad en el ciclo vital. Volumen I. 2016. 167–172 p. Available from: [http://formacionasunivep.com/IIIcongresosalud/documents/Libro\\_Salud\\_alimentacion\\_y\\_sexualidad\\_en\\_el\\_ciclo\\_vital\\_Volumen\\_I.pdf#page=168](http://formacionasunivep.com/IIIcongresosalud/documents/Libro_Salud_alimentacion_y_sexualidad_en_el_ciclo_vital_Volumen_I.pdf#page=168)
161. Paino Pardal L, Poblet i Montells L, Ríos Álvarez L. Mayores que viven solos y malnutrición. Estudio SOLGER. Atención Primaria [Internet]. 2017 Oct [cited 2018 Sep 7];49(8):450–8. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0212656716305637>
162. Cutillas AB, Herrero E, de San Eustaquio A, Zamora S, Pérez-Llamas F. Prevalencia de peso insuficiente, sobrepeso y obesidad, ingesta de energía y perfil calórico de la dieta de estudiantes universitarios de la comunidad autónoma de la región de Murcia (España). Nutr Hosp. 2013;28(3):683–9.
163. Careau V. Energy Intake, Basal Metabolic Rate, and Within-Individual Trade-Offs in Men and Women Training for a Half Marathon: A Reanalysis. Physiol Biochem Zool. 2017;90(3):392–8.
164. Srivastava R, Batra A, Dhawan D, Bakhshi S. Association of energy intake and expenditure with obesity: A cross-sectional study of 150 pediatric patients following treatment for leukemia. Pediatr Hematol Oncol. 2017;34(1):29–35.

165. Bonfanti N, Fernandez JM, Gomez-Delgado F, Perez-Jimenez F. [Effect of two hypocaloric diets and their combination with physical exercise on basal metabolic rate and body composition]. *Nutr Hosp.* 2014 Mar;29(3):635–43.
166. Farias G, Thieme RD, Teixeira LM, Heyde ME, Bettini S, Radominski R. *Nutrición Hospitalaria Trabajo Original.* *Nutr Hosp.* 2016;33(5):1108–15.
167. Zaragoza Martí A, Ferrer Cascales R, Cabañero Martínez MJ, Hurtado Sánchez JA, Laguna Pérez A. Adherencia a la dieta mediterránea y su relación con el estado nutricional en personas mayores. *Nutr Hosp.* 2015;31(4):1667–74.
168. Abreu-Reyes JA, Álvarez-Luis D, Arteaga-Hernández V, Sánchez-Mendez M, Abreu-González R. Mediterranean diet adherence by patients with primary open angle glaucoma. *Arch la Soc Española Oftalmol (English Ed.* 2017;92(8):353–8.
169. Romero Pérez A, Rivas Velasco A. Adherence to Mediterranean diet and bone health. *Nutr Hosp.* 2014;29(5):989–96.
170. Savanelli MC, Barrea L, Macchia PE, Savastano S, Falco A, Renzullo A, et al. Preliminary results demonstrating the impact of Mediterranean diet on bone health. *J Transl Med.* 2017 Apr;15(1):81.
171. Melaku YA, Gill TK, Taylor AW, Adams R, Shi Z. Association between nutrient patterns and bone mineral density among ageing adults. *Clin Nutr ESPEN.* 2017;
172. Stefan L, Cule M, Milinovic I, Sporis G, Juranko D. The relationship between adherence to the Mediterranean diet and body composition in Croatian university students. *Eur J Integr Med.* 2017;13(Supplement C):41–6.
173. Marventano S, Godos J, Platania A, Galvano F, Mistretta A, Grosso G. Mediterranean diet adherence in the Mediterranean healthy eating, aging and lifestyle (MEAL) study cohort. *Int J Food Sci Nutr.* 2017 May;1–8.
174. Grao-Cruces A, Nuviala A, Fernandez-Martinez A, Martinez-Lopez E-J. Relationship of physical activity and sedentarism with tobacco and alcohol consumption, and Mediterranean diet in Spanish teenagers. *Nutr Hosp.* 2015 Apr;31(4):1693–700.



# DIFUSIÓN DE RESULTADOS

## 1. COMUNICACIONES A CONGRESOS:

- **Título: Machine Learning para la identificación de factores predictivos del grado de adherencia a la DM.** Autores: A. Pazos, A. Arceo Vilas, C. Fernández Lozano, S. Álvarez González, S. Pérttega Díaz. XXII Congreso Nacional de Informática de la Salud. INFORSALUD. Track. Madrid. 5-7 marzo 2019.
- **Título: Medición de la grasa corporal mediante ecuaciones antropométricas e impedancia bioeléctrica.** Autores: Arceo Vilas, A., López Calviño, B., Piñón Villar, C., García-Rodríguez, MT.; Pita Fernández, S., Pérttega Díaz, S., Seoane Pillado, T. XVIII Encuentro Internacional de Investigación en Cuidados. Victoria-Gasteiz. 11-14 Noviembre 2014
- **Título: Insatisfacción corporal y obsesión por la delgadez en estudiantes de enfermería de la Universidad de La Coruña.** Autores: García-Rodríguez, MT., Pita Fernández, S., López Calviño, B., Patiño Vázquez, A., Pérttega Díaz, S., Seoane Pillado, T., Piñón Villar, C., Arceo-Vilas, A Rodríguez Camacho, E. XVIII Encuentro Internacional de Investigación en Cuidados. Victoria- Gasteiz. 11-14 Noviembre 2014
- **Título: La dinamometría y la calidad de vida de los pacientes en lista de trasplante hepático según el estado nutricional.** Autores: García-Rodríguez, MT.; Piñón Villar, MC., López Calviño, B., Pita Fernández, S., Arceo Vilas. A. XVIII Encuentro Internacional de Investigación en Cuidados. Victoria- Gasteiz. 11-14 Noviembre 2014.
- **Título: Prevalencia de deterioro cognitivo y dependencia para actividades básicas e instrumentales de la vida diaria.** Autores: Pita Fernández, S., López Calviño, B., Pérttega Díaz, S., Seoane Pillado, T., García Alonso, P., Arceo Vilas, A. XXV Congreso Internacional Sociedad Gallega de Gerontología y Geriatria. A Coruña. 10-11 mayo 2013. Premio a la mejor Comunicación oral clínica presentada.
- **Título: Valoración del estado nutricional en población de 40 y más años.** Autores: Arceo-Vilas, A. Seoane Pillado, T., Pita Fernández, S., López Calviño, B., Pérttega Díaz, S., García Cajide, Y., García-Rodríguez, MT. XVII Encuentro Internacional de Investigación en Cuidados. Lleida. 12-15 Noviembre 2013.

## 2. PUBLICACIONES EN REVISTAS INDEXADAS:

- Arceo-Vilas A, Fernandez-Lozano C, Pita S, Pérttega-Díaz S, Pazos A. 2020. Identification of predictive factors of the degree of adherence to the Mediterranean diet through machine-learning techniques. PeerJ Computer Science 6:e287 <https://doi.org/10.7717/peerj-cs.287>.

# ANEXOS

**TÍTULO DEL ESTUDIO:**

Valoración del estado nutricional y adherencia a la dieta mediterránea en una muestra aleatoria poblacional de 40 y más años en el municipio de Cambre.

**INVESTIGADORA PRINCIPAL:**

Alba Arceo Vilas. Unidad de Cuidados Intensivos.

Complejo Hospitalario Universitario A Coruña.

**CENTRO:**

Centro de Salud de Cambre.

**INTRODUCCIÓN:**

Este documento tiene como objetivo informarle sobre un estudio de investigación en el que se le invita a participar. Nuestra intención es tan solo que usted reciba la información correcta y suficiente para que pueda evaluar y juzgar si quiere o no participar en este estudio. Para ello lea esta hoja informativa con atención y nosotros le aclararemos las dudas que le puedan surgir después de la explicación.

**PARTICIPACIÓN VOLUNTARIA:**

Debe saber que su participación en este estudio es completamente voluntaria y que puede decidir no participar o cambiar su decisión y retirar el consentimiento en cualquier momento, sin que por ello se altere la relación con su médico ni se produzca perjuicio alguno en la asistencia sanitaria a la que usted tiene derecho.

**OBJETIVO:**

El objetivo de este estudio es determinar el estado nutricional y la adherencia a la dieta mediterránea de la población de 40 y más años del municipio de Cambre.

Necesitaremos conocer su peso, talla, perímetro de cintura, perímetro de cadera, perímetro de brazo, mediremos 4 pliegues cutáneos y obtendremos el porcentaje de grasa corporal a través de una balanza.

También se le harán unas preguntas sobre si le han sido diagnosticadas algunas enfermedades. Finalmente, se le pedirá que conteste a unos cuestionarios en los que se le preguntará sobre diferentes aspectos relacionados con nutrición y la dieta

#### **BENEFICIOS OBTENIDOS DE SU PARTICIPACIÓN:**

La realización de este estudio no va a tener ningún efecto sobre las pruebas o consultas que se le realicen en su seguimiento. Es probable que de una manera inmediata este estudio no tenga ningún beneficio sobre usted, pero ayudará a determinar los factores que se asocian al estado nutricional y autopercepción corporal y su grado de control actual. Asimismo, el determinar las características de la población de 40 y más años ayudará a conocer más a esta población y sus condiciones de salud.

#### **RIESGOS E INCONVENIENTES:**

La realización de este estudio no conlleva un aumento ni el número de consultas, ni el número de análisis, ni el número de pruebas de ningún tipo (radiografías, ecografías) que se le puedan realizar.

#### **UTILIZACIÓN DE LA INFORMACIÓN OBTENIDA SOBRE MÍ:**

La información no le identificará por su nombre. La información será sólo identificada por un número de código con el objeto de proteger su privacidad. Asimismo, si los resultados del estudio se publican su identidad permanecerá confidencial. Solamente la enfermera encargada del estudio tendrá una lista que relacione el número del código con su nombre. Con la firma de este consentimiento informado, usted autoriza el acceso directo a su historial clínico para el estudio de sus enfermedades.

La información recogida sobre usted será confidencial indefinidamente. El tratamiento, comunicación y cesión de sus datos se hará conforme a lo dispuesto por la Ley Orgánica 15/1999, de 13 de Diciembre, de Protección de Datos de Carácter Personal. En todo momento, usted podrá acceder a sus datos, corregirlos o cancelarlos.

#### **INTERESES ECONÓMICOS:**

Usted no percibirá remuneración alguna por participar en el estudio. Ninguno de los investigadores participantes en el proyecto recibirá tampoco retribuciones específicas por su dedicación en el estudio. **Muchas gracias por su colaboración.**

*Trabajo de Investigación en*  
**C.AMBRE**

La cardiopatía isquémica (CI) es uno de los principales problemas de salud pública en el momento actual, además de la primera causa de mortalidad en la población adulta de los países desarrollados.

La enfermedad pulmonar obstructiva crónica (EPOC) es una enfermedad potencialmente mortal, prevenible pero incurable, su análisis epidemiológico es difícil debido a las dificultades conceptuales y de diagnóstico que conlleva, la EPOC suele empeorar con el tiempo por lo que generalmente se diagnostica en mayores de 40 años. Por todo ello es necesario realizar estudios que revelen su impacto real.

Los cuestionarios de salud autopercibida permiten medir la parte subjetiva de la salud y conocer la evaluación que hace el individuo de su estado de bienestar general, esto proporciona información complementaria que puede ser muy útil en el manejo y evaluación de los pacientes.

La prevención en geriatría permite reducir la mortalidad prematura originada por enfermedades agudas y crónicas, mantener la independencia funcional de la persona tanto como sea posible, aumentar la esperanza de vida activa (independiente) y mejorar la calidad de vida.

Una apropiada salud del pie repercute en la salud general del individuo, se pueden prevenir muchos problemas importantes del pie mediante una atención apropiada.



**COMPLEJO HOSPITALARIO  
UNIVERSITARIO A CORUÑA**

**Y**

**CENTRO DE SALUD DE CAMBRE**

*Teléfono de contacto 683532664*

*Aprobado por el CEIC de Galicia  
cód. reg. 2008/264*

*Premio Dr. Jiménez Herrero, 2011  
Otorgado por la Real Academia de Medicina y Cirugía de Galicia*

*Trabajo de Investigación en*  
**C.AMBRE**

**RIESGO CARDIOVASCULAR,  
COMORBILIDAD, CALIDAD  
DE VIDA Y DEPENDENCIA  
EN PERSONAS DE 40 Y MÁS  
AÑOS DE EDAD EN EL  
MUNICIPIO DE CAMBRE**



**CENTRO DE SALUD DE CAMBRE  
COMPLEJO HOSPITALARIO  
UNIVERSITARIO DE A CORUÑA**

## ANEXO 2. TRÍPTICO INFORMATIVO DEL ESTUDIO (HOJA 2)

### *Trabajo de investigación en el Municipio de Cambre*

Este documento tiene como objetivo informarle sobre un estudio de investigación en el que se le invita a participar. Nuestra intención es tan solo que usted reciba la información correcta y suficiente para que pueda evaluar y juzgar si quiere o no participar en este estudio. Para ello lea esta hoja informativa con atención y nosotros le aclararemos las dudas que le puedan surgir después de la explicación.

**PARTICIPACIÓN VOLUNTARIA:** Debe saber que su participación en este estudio es completamente voluntaria y que puede decidir no participar o cambiar su decisión y retirar el consentimiento en cualquier momento, sin que por ello se altere la relación con su médico ni se produzca perjuicio alguno en la asistencia sanitaria a la que usted tiene derecho. Usted no percibirá remuneración alguna por participar en el estudio. Ninguno de los investigadores recibirán tampoco retribuciones específicas por su dedicación en el estudio.

**OBJETIVO:** El objetivo de este estudio es determinar el riesgo cardiovascular, la comorbilidad (coexistencia de otras enfermedades), la calidad de vida y el nivel de dependencia de la población de 40 y más años del municipio de Cambre.

Para determinar el riesgo cardiovascular necesitamos conocer su tensión arterial, si es usted o no diabético, su colesterol en sangre, si fuma o no y si el tamaño de su corazón está agrandado tras realizarle un electrocardiograma. Para determinar los factores de riesgo cardiovascular se le va a pesar, tallar, tomar su tensión arterial y se le realizará un análisis de sangre para determinar los niveles de lípidos (grasas), colesterol y nivel de azúcar. Dicho análisis no será necesario si tiene ya uno reciente (en los últimos 12 meses). También se le harán unas preguntas acerca de su consumo de tabaco, sobre si le han sido diagnosticadas algunas enfermedades y la medicación que consume habitualmente.

Se le medirá la concentración de oxígeno en sangre mediante un pequeño aparato (pulsioxímetro) que se le colocará en un dedo. Además, se evaluará la capacidad de sus pulmones mediante una prueba sencilla en la que en varias ocasiones se le pedirá que coja la máxima cantidad de aire posible en sus pulmones para expulsarla a continuación. Se le realizará además un examen de sus pies por personal cualificado. Ninguna de estas pruebas supondrá la realización de analíticas, radiografías, ecografías u otras técnicas invasivas.

Si tiene entre 45 y 64 años, se le pedirá que conteste a unos cuestionarios en los que se le preguntará sobre diferentes aspectos relacionados con su calidad de vida, niveles de ansiedad y la frecuencia con la que realiza ejercicio físico.

Si tiene usted más de 65 años, también se evaluará la presencia de deterioro cognitivo (alteraciones en la memoria) realizándole algunas preguntas. Finalmente, se le pedirá que conteste a unos cuestionarios en los que se le preguntará sobre diferentes aspectos relacionados con su calidad de vida, disponibilidad de recursos sociales y su capacidad para realizar diferentes actividades en la vida diaria.

**BENEFICIOS OBTENIDOS DE SU PARTICIPACIÓN:** La realización de este estudio no va a tener ningún efecto sobre las pruebas, análisis o consultas que se le realicen en su seguimiento. Es probable que de una manera inmediata este estudio no tenga ningún beneficio sobre usted, pero ayudará a determinar los factores que se asocian a la presencia de enfermedades cardiovasculares y su grado de control actual.

**RIESGOS E INCONVENIENTES:** La realización de este estudio no conlleva un aumento ni el número de consultas, ni el número de análisis, ni el número de pruebas de ningún tipo (radiografías, ecografías) que se le puedan realizar. Incluso, si su médico le ha realizado un análisis reciente (en los últimos 12 meses) no será necesario hacerle ningún análisis más. La realización del análisis de sangre puede producir, en la zona del pinchazo, ocasionalmente un pequeño hematoma.

**UTILIZACIÓN DE LA INFORMACIÓN:** La información no le identificará por su nombre, sólo identificada por un número de código con el objeto de proteger su privacidad. Asimismo, si los resultados del estudio se publican su identidad permanecerá confidencial. Solamente los médicos y enfermeras encargadas del estudio tendrán una lista que relacione el número del código con su nombre. Con la firma de este consentimiento informado, usted autoriza el acceso directo a su historial clínico para el estudio de sus enfermedades.

La información recogida sobre usted será confidencial indefinidamente. El tratamiento, comunicación y cesión de sus datos se hará conforme a lo dispuesto por la Ley Orgánica 15/1999, de 13 de Diciembre, de Protección de Datos de Carácter Personal. En todo momento, usted podrá acceder a sus datos, corregirlos o cancelarlos.

ANEXO 3. CONSENTIMIENTO INFORMADO DEL PACIENTE

Yo, Don/Dña..... (EN MAYÚSCULAS), con DNI nº....., confirmo que:

- He sido informado/a verbalmente y por escrito sobre el estudio “RIESGO CARDIOVASCULAR, COMORBILIDAD, CALIDAD DE VIDA Y DEPENDENCIA EN PERSONAS DE 40 Y MÁS AÑOS DE EDAD EN EL MUNICIPIO DE CAMBRE” por D./Dña. .... (EN MAYÚSCULAS)

- Entiendo el propósito del estudio y he tenido oportunidad de realizar preguntas sobre el mismo.

- Entiendo que la participación en este proyecto es voluntaria y que soy libre de abandonar el estudio en cualquier momento sin dar ninguna razón y sin que mis derechos de atención médica o legal se vean afectados.

- Entiendo que la participación en este proyecto no interferirá sobre el seguimiento y/o el tratamiento que pueda recibir por otros problemas.

Y estoy de acuerdo en participar en este estudio, lo cual implica que:

- Doy mi consentimiento para realizar las preguntas anteriormente descritas, las pruebas detalladas y la consulta de los datos de mi historia clínica.

Firma del/la paciente

Firma del informador/a

En....., a..... de ..... de 201...

## ANEXO 4. CUADERNO DE RECOGIDA DE DATOS (40-64 AÑOS)

### CUADERNO DE RECOGIDA DE DATOS. POBLACIÓN ADULTA 40 Y MÁS AÑOS EN EL MUNICIPIO DE CAMBRE

<b>Paciente</b>
Nº historia: _____
Nombre: _____
Apellidos: _____
Firma entrevistador: _____

#### Variables identificación del paciente:

Fecha nacimiento (dd/mm/aaaa):

Sexo:

Hombre

Mujer

Lugar:

Teléfono:

Nivel de estudios:  sin estudios

1º grado incompleto

1º grado (Bachillerato Elemental, EGB)

2º grado (Bachillerato Superior, BUP, FP)

Universitario

#### Variables antropométricas:

Peso:

IMC:

Talla:

Perímetro de cintura:

Perímetro de cadera:

Actividad física:

Pliegues cutáneos:

Tricipital:

Subescapular:

Bicipital:

Suprailíaco:

Perímetro del brazo:

Perímetro de la pantorrilla:

#### **Bioimpedancia:**

Masa adiposa:

Masa ósea (kg):

Masa líquida:

MTB de actividad en Kcal:

Masa muscular:

MTB basal calórico en Kcal:

**CUADERNO DE RECOGIDA DE  
DATOS. POBLACIÓN ADULTA 40  
Y MÁS AÑOS EN EL MUNICIPIO  
DE CAMBRE**

<b>Paciente</b>
Nº historia: _____
Nombre: _____
Apellidos: _____
<b>Firma entrevistador:</b>

**Variables riesgo cardiovascular:**

Fumador:

- no
- ex-fumador (el paciente dejó de fumar más de 12 meses antes de entrar en el estudio)
- si (el paciente fuma o ha fumado en los 12 últimos meses, al menos 1 cigarrillo o 1 pipa o 1 puro al día). Test de Fagerström.

TAS/TAD:

	TAD	TAS
Medición 1		
Medición 2		

Analítica:

Previa: fecha \_\_\_\_\_ no

Actual: fecha \_\_\_\_\_ no **(pedirla)**

EKG:

Hipertrofia ventrículo izquierdo (criterio de Sokolow):

- si  no

Otros hallazgos:

- FA
- Bloqueo AV:
- Extrasistolias
- Isquemia
- Necrosis

**CUADERNO DE RECOGIDA DE  
DATOS. POBLACIÓN ADULTA 40  
Y MÁS AÑOS EN EL MUNICIPIO  
DE CAMBRE**

<b>Paciente</b>
Nº historia: _____
Nombre: _____
Apellidos: _____
<b>Firma entrevistador:</b>

Enfermedades:

ENFERMEDAD	HISTORIA	TEST CHARLSON
Infarto miocárdico		
Insuficiencia cardíaca		
Enfermedad vascular periférica		
Enfermedad vascular cerebral		
Demencia		
Hemiplejía		
Enfermedad pulmonar crónica		
Enfermedad tejido conectivo		
Úlcera péptica		
Hepatopatía leve		
Hepatopatía moderada-severa (ascitis y cirrosis)		
Diabetes mellitus		
Diabetes con afectación orgánica		
Insuficiencia renal moderada o grave		
Cáncer, leucemia, linfoma		
Cáncer con metástasis		
SIDA		
HTA		
Alteración de lípidos		
Otras		

Fecha en que se realiza la entrevista (dd/mm/aaaa):

## ANEXO 5. CUADERNO DE RECOGIDA DE DATOS (65 Y MÁS AÑOS)

### CUADERNO DE RECOGIDA DE DATOS. POBLACIÓN ADULTA 65 Y MÁS AÑOS EN EL MUNICIPIO DE CAMBRE

<b>Paciente</b>
Nº historia: _____
Nombre: _____
Apellidos: _____
Firma entrevistador: _____

#### Variables identificación del paciente

Fecha nacimiento (dd/mm/aaaa):

Sexo:

Hombre

Mujer

Lugar:

Teléfono:

Nivel de estudios:  sin estudios

1º grado incompleto

1º grado (Bachillerato Elemental, EGB)

2º grado (Bachillerato Superior, BUP, FP)

Universitario

#### Variables antropométricas:

Peso:

IMC:

Talla:

Perímetro de cintura:

Perímetro de cadera:

Pliegues cutáneos:

Tricipital:

Subescapular:

Bicipital:

Suprailíaco:

Perímetro del brazo:

Perímetro de la pantorrilla:

**Bioimpedancia:** Grado de actividad:

Masa adiposa:

Masa ósea (kg):

Masa líquida:

MTB de actividad en Kcal:

Masa muscular:

MTB basal calórico en Kcal:

**CUADERNO DE RECOGIDA DE  
DATOS. POBLACIÓN ADULTA 65 Y  
MÁS AÑOS EN EL MUNICIPIO DE  
CAMBRE**

<b>Paciente</b>
Nº historia: _____
Nombre: _____
Apellidos: _____
 Firma entrevistador: _____

**Variables riesgo cardiovascular:**

Fumador:

- no
- ex-fumador (el paciente dejó de fumar más de 12 meses antes de entrar en el estudio)
- si (el paciente fuma o ha fumado en los 12 últimos meses, al menos 1 cigarrillo o 1 pipa o 1 puro al día). Test de Fagerström.

TAS/TAD:

	TAD	TAS
Medición 1		
Medición 2		

Analítica:

- Previa: fecha \_\_\_\_\_ no
- Actual: fecha \_\_\_\_\_ no **(pedirla)**

Actividad física (más de 30 minutos de actividad física moderada-intensa más de 5 días a la semana):

- si  no

EKG:

Hipertrofia ventrículo izquierdo (criterio de Sokolow):

- si  no

Otros hallazgos:

- FA
- Bloqueo AV:
- Extrasistolias
- Isquemia
- Necrosis

**CUADERNO DE RECOGIDA DE  
DATOS. POBLACIÓN ADULTA 65 Y  
MÁS AÑOS EN EL MUNICIPIO DE  
CAMBRE**

<b>Paciente</b>
Nº historia: _____
Nombre: _____
Apellidos: _____
<b>Firma entrevistador:</b>

Enfermedades:

ENFERMEDAD	HISTORIA	TEST CHARLSON
Infarto miocárdico		
Insuficiencia cardíaca		
Enfermedad vascular periférica		
Enfermedad vascular cerebral		
Demencia		
Hemiplejia		
Enfermedad pulmonar crónica		
Enfermedad tejido conectivo		
Úlcera péptica		
Hepatopatía leve		
Hepatopatía moderada-severa (ascitis y cirrosis)		
Diabetes mellitus		
Diabetes con afectación orgánica		
Insuficiencia renal moderada o grave		
Cáncer, leucemia, linfoma		
Cáncer con metástasis		
SIDA		
HTA		
Alteración de lípidos		
Otras		

Fecha en que se realiza la entrevista (dd/mm/aaaa):

**ANEXO 6. SCORE DE COMORBILIDAD DE CHARLSON**

**ESCALA DE COMORBILIDAD DE CHARLSON**

<b>Paciente</b> N° historia: _____ Nombre: _____ Apellidos: _____  Firma entrevistador: _____
--

Edad del paciente: \_\_\_\_ años

		<b>Sí</b>	<b>No</b>	
SIDA (no únicamente HIV positivo)		<input type="checkbox"/>	<input type="checkbox"/>	
Enfermedad Cerebrovascular		<input type="checkbox"/>	<input type="checkbox"/>	
EPOC		<input type="checkbox"/>	<input type="checkbox"/>	
Insuficiencia Cardíaca Congestiva		<input type="checkbox"/>	<input type="checkbox"/>	
Enfermedad de tejido conectivo		<input type="checkbox"/>	<input type="checkbox"/>	
Demencia		<input type="checkbox"/>	<input type="checkbox"/>	
Hemiplejía		<input type="checkbox"/>	<input type="checkbox"/>	
Leucemia (aguda o crónica)		<input type="checkbox"/>	<input type="checkbox"/>	
Linfoma		<input type="checkbox"/>	<input type="checkbox"/>	
Infarto de Miocardio		<input type="checkbox"/>	<input type="checkbox"/>	
Enfermedad Vascul ar Periférica (Incluye aneurisma de aorta 6cm)		<input type="checkbox"/>	<input type="checkbox"/>	
Úlcus péptico		<input type="checkbox"/>	<input type="checkbox"/>	
	<b>No</b>	<input type="checkbox"/> <b>Sin afectación</b>	<input type="checkbox"/> <b>Con afectación</b>	
		<input type="checkbox"/> <b>de órganos</b>	<input type="checkbox"/> <b>de órganos</b>	
Diabetes Mellitus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<b>No</b>	<b>Leve</b>	<b>Moderada</b>	<b>Severa</b>
Enfermedad Hepática	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enfermedad renal	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<b>No</b>	<b>Sin metástasis</b>	<b>Con metástasis</b>	
Tumor sólido maligno	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

## TEST DE FAGERSTRÖM

<b>Paciente</b>
Nº historia: _____
Nombre: _____
Apellidos: _____
<b>Firma entrevistador:</b>

El test de Fagerström mide el grado de adicción al tabaco. Consta de 6 preguntas:

1. ¿Cuántos minutos pasan entre el momento de levantarse y fumar el primer cigarrillo?

- 5 ó menos
- de 6 a 30
- de 31 a 60
- más de 60

2. ¿Encuentra dificultad para abstenerse de fumar en lugares donde está prohibido?

(cines, hospitales, metro...)

- Sí
- No

2. ¿Cuál es el cigarrillo que más necesita?

- El primero de la mañana
- Otro

4. ¿Cuántos cigarrillos fuma al día?

- más de 30
- entre 21 y 30
- entre 11 y 20
- menos de 11

5. ¿Fuma más durante las primeras horas de la mañana que durante el resto del día?

- Sí
- No

6. ¿Fuma usted aunque esté tan enfermo como para estar en la cama?

Sí

No

## **EVALUACIÓN DE LOS ESTADIOS DE CAMBIO**

1. En el último año, ¿Cuántas veces dejó de fumar durante al menos 24 horas?

Ninguna

Una o más veces

2. Actualmente fumo, pero tengo la intención de dejarlo dentro de los próximos 30 días

Sí

No

3. Actualmente fumo, pero tengo la intención de dejarlo dentro de los próximos 6 meses

Sí

No

4. Actualmente fumo, pero NO tengo la intención de dejarlo dentro de los próximos 6 meses

Sí

No

### **Resultado:**

*Precontemplación: Respuesta afirmativa a la pregunta 4*

*Contemplación: Respuesta afirmativa a la pregunta 3*

*Preparación: Respuestas afirmativas a las preguntas 2 y a la 1 (una o más veces)*

## ANEXO 8. CUESTIONARIO MNA (MINI NUTRITIONAL ASSESSMENT)

Apellidos:		Nombre:		
Sexo:	Edad:	Peso, kg:	Altura, cm:	Fecha:

Responda a la primera parte del cuestionario indicando la puntuación adecuada para cada pregunta. Sume los puntos correspondientes al cribaje y si la suma es igual o inferior a 11, complete el cuestionario para obtener una apreciación precisa del estado nutricional.

Cribaje	
<b>A</b> Ha perdido el apetito? Ha comido menos por faltado apetito, problemas digestivos, dificultades de masticación o deglución en los últimos 3 meses? 0 = ha comido mucho menos 1 = ha comido menos 2 = ha comido igual	<input type="checkbox"/>
<b>B</b> Pérdida reciente de peso (<3 meses) 0 = pérdida de peso > 3 kg 1 = no lo sabe 2 = pérdida de peso entre 1 y 3 kg 3 = no ha habido pérdida de peso	<input type="checkbox"/>
<b>C</b> Movilidad 0 = de la cama al sillón 1 = autonomía en el interior 2 = sale del domicilio	<input type="checkbox"/>
<b>D</b> Ha tenido una enfermedad aguda o situación de estrés psicológico en los últimos 3 meses? 0 = sí    2 = no	<input type="checkbox"/>
<b>E</b> Problemas neuropsicológicos 0 = demencia o depresión grave 1 = demencia moderada 2 = sin problemas psicológicos	<input type="checkbox"/>
<b>F</b> Índice de masa corporal (IMC = peso / (talla) <sup>2</sup> en kg/m <sup>2</sup> ) 0 = IMC < 19 1 = 19 ≤ IMC < 21 2 = 21 ≤ IMC < 23. 3 = IMC ≥ 23.	<input type="checkbox"/>
<b>Evaluación del cribaje</b> (subtotal máx. 14 puntos)	<input type="checkbox"/> <input type="checkbox"/>
12-14 puntos:                      estado nutricional normal 8-11 puntos:                        riesgo de malnutrición 0-7 puntos:                          malnutrición	
Para una evaluación más detallada, continúe con las preguntas G-R	
Evaluación	
<b>G</b> El paciente vive independiente en su domicilio? 1 = sí    0 = no	<input type="checkbox"/>
<b>H</b> Toma más de 3 medicamentos al día? 0 = sí    1 = no	<input type="checkbox"/>
<b>I</b> Úlceras o lesiones cutáneas? 0 = sí    1 = no	<input type="checkbox"/>
<b>J</b> Cuántas comidas completas toma al día? 0 = 1 comida 1 = 2 comidas 2 = 3 comidas	<input type="checkbox"/>
<b>K</b> Consume el paciente <ul style="list-style-type: none"> <li>• productos lácteos al menos una vez al día?                      sí <input type="checkbox"/> no <input type="checkbox"/></li> <li>• huevos o legumbres 1 o 2 veces a la semana?                      sí <input type="checkbox"/> no <input type="checkbox"/></li> <li>• carne, pescado o aves, diariamente?                      sí <input type="checkbox"/> no <input type="checkbox"/></li> </ul> 0.0 = 0 o 1 síes 0.5 = 2 síes 1.0 = 3 síes	<input type="checkbox"/> <input type="checkbox"/>
<b>L</b> Consume frutas o verduras al menos 2 veces al día? 0 = no    1 = sí	<input type="checkbox"/>
<b>M</b> Cuántos vasos de agua u otros líquidos toma al día? (agua, zumo, café, té, leche, vino, cerveza...) 0.0 = menos de 3 vasos 0.5 = de 3 a 5 vasos 1.0 = más de 5 vasos	<input type="checkbox"/> <input type="checkbox"/>
<b>N</b> Forma de alimentarse 0 = necesita ayuda 1 = se alimenta solo con dificultad 2 = se alimenta solo sin dificultad	<input type="checkbox"/>
<b>O</b> Se considera el paciente que está bien nutrido? 0 = malnutrición grave 1 = no lo sabe o malnutrición moderada 2 = sin problemas de nutrición	<input type="checkbox"/>
<b>P</b> En comparación con las personas de su edad, cómo encuentra el paciente su estado de salud? 0.0 = peor 0.5 = no lo sabe 1.0 = igual 2.0 = mejor	<input type="checkbox"/> <input type="checkbox"/>
<b>Q</b> Circunferencia braquial (CB en cm) 0.0 = CB < 21 0.5 = 21 ≤ CB ≤ 22 1.0 = CB > 22	<input type="checkbox"/> <input type="checkbox"/>
<b>R</b> Circunferencia de la pantorrilla (CP en cm) 0 = CP < 31 1 = CP ≥ 31	<input type="checkbox"/>
<b>Evaluación</b> (máx. 16 puntos)	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<b>Cribaje</b>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<b>Evaluación global</b> (máx. 30 puntos)	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
Evaluación del estado nutricional	
De 24 a 30 puntos: <input type="checkbox"/> estado nutricional normal De 17 a 23.5 puntos: <input type="checkbox"/> riesgo de malnutrición Menos de 17 puntos: <input type="checkbox"/> malnutrición	

Ref: Vellas B, Villars H, Abellan G, et al. Overview of the MNA® - its History and Challenges. J Nut Health Aging 2006 ; 10 : 456-465.  
 Rubenstein LZ, Harker JO, Salva A, Gulgoz Y, Vellas B. Screening for Undernutrition in Geriatric Practice : Developing the Short-Form Mini Nutritional Assessment (MNA-SF). J Geront 2001 ; 56A : M366-377.  
 Gulgoz Y. The Mini-Nutritional Assessment (MNA®) Review of the Literature - What does it tell us? J Nutr Health Aging 2006 ; 10 : 466-467.  
 © Société des Produits Nestlé, S.A., Vevey, Switzerland, Trademark Owners  
 © Nestlé, 1994, Revision 2006. N67200 12/99 10M  
 Para más información: [www.mna-elderly.com](http://www.mna-elderly.com)

## ANEXO 9. TEST DE ADHERENCIA A LA DIETA MEDITERRÁNEA

<b>Paciente</b>
Nº historia: _____
Nombre: _____
Apellidos: _____
<b>Firma entrevistador:</b>

### CUESTIONARIO DE ADHERENCIA A LA DIETA MEDITERRÁNEA

Nos interesa conocer sus hábitos de alimentación, por ello necesitamos que responda a algunas cuestiones relacionadas con su dieta. Le informamos que este cuestionario es totalmente anónimo.

Muchas gracias por su colaboración

Nº	Pregunta	Modo de valoración	Puntos
1	¿Usa usted el aceite de oliva principalmente para cocinar?	Si =1 punto	<input type="checkbox"/>
2	¿Cuánto aceite de oliva consume en total al día (incluyendo el usado para freír, el de las comidas fuera de casa, las ensaladas, etc.)?	Dos o más cucharadas =1 punto	<input type="checkbox"/>
3	¿Cuántas raciones de verdura u hortalizas consume al día (las guarniciones acompañamientos contabilizan como ½ ración)?	Dos o más al día (al menos una de ellas en ensaladas o crudas)= 1 punto	<input type="checkbox"/>
4	¿Cuántas piezas de fruta (incluyendo zumo natural) consume al día?	Tres o más al día= 1 punto	<input type="checkbox"/>
5	¿Cuántas raciones de carnes rojas, hamburguesas, salchichas o embutidos consume al día (una ración equivale a 100-150 gr.)?	Menos de una al día= 1 punto	<input type="checkbox"/>
6	¿Cuántas raciones de mantequilla, margarina o nata consume al día porción individual equivale a 12 gr)?	Menos de una al día= 1 punto	<input type="checkbox"/>
7	¿Cuántas bebidas carbonatadas y/o azucaradas (refrescos, colas, tónicas, bitter) consume al día?	Menos de una al día= 1 punto	<input type="checkbox"/>
8	¿Bebe vino? ¿Cuánto consume a la semana?	Tres o más vasos por semana= 1 punto	<input type="checkbox"/>
9	¿Cuántas raciones de legumbres consume a la semana (una ración o plato equivale a 150 gr)?	Tres o más por semana= 1 punto	<input type="checkbox"/>
10	¿Cuántas raciones de pescado o mariscos consume a la semana (un plato, pieza o ración equivale a 100-150 gr de pescado ó 4-5 piezas de marisco)?	Tres o más por semana= 1 punto	<input type="checkbox"/>
11	¿Cuántas veces consume repostería comercial (no casera) como galletas, flanes, dulces o pasteles a la semana?	Menos de tres por semana= 1 punto	<input type="checkbox"/>
12	¿Cuántas veces consume frutos secos a la semana (una ración equivale a 30 gr)?	Una o más por semana= 1 punto	<input type="checkbox"/>
13	¿Consume preferentemente carne de pollo, pavo o conejo en vez de ternera, cerdo, hamburguesas o salchichas (carne de pollo: una pieza o ración equivale a 100- 150 gr)?	Si= 1 punto	<input type="checkbox"/>
14	¿Cuántas veces a la semana consume los vegetales cocinados, la pasta, el arroz u otros platos aderezados con una salsa de tomate, ajo, cebolla o puerro elaborada a fuego lento con aceite de oliva (sofrito)?	Dos o más por semana= 1 punto	<input type="checkbox"/>
Resultado final, puntuación total			<input type="checkbox"/>

**PUNTUACIÓN TOTAL:**  
 < 9 baja adherencia  
 >= 9 buena adherencia

ANEXO 10. CUESTIONARIO EATING DISORDER INVENTORY (EDI-IC Y EDI-OD)

<b>Paciente</b>
Nº historia: _____
Nombre: _____
Apellidos: _____
<b>Firma entrevistador:</b>

**Percepción subjetiva de peso:**

Considero que mi peso es:

- A) superior a lo normal
- B) normal
- C) inferior a lo normal

**Subescala de Eating Disorder Inventory (EDI-IC):  
insatisfacción corporal**

Instrucciones:

A continuación, encontrará unas frases que se utilizan corrientemente para medir la insatisfacción del paciente con la forma general de su cuerpo o con la de aquellas partes del mismo que más preocupan a los que padecen trastornos de la conducta alimentaria (estómago, caderas, muslos, nalgas...). Marque con una X lo que mejor indique como se siente ahora mismo.

**Subescala de Eating Disorder Inventory (EDI-O): obsesión  
por la delgadez**

Instrucciones:

Los elementos de esta escala se refieren a la preocupación por el peso, las dietas y el miedo a engordar. Marque con una X lo que mejor indique como se siente ahora mismo.

SUBESC.EDI		NUNCA	POCAS VECES	A VECES	A MENUDO	CASI SIEMPRE	SIEMPRE
IC 1	Creo que mi estómago es demasiado grande	0	0	0	1	2	3
O 1	Como dulces e hidratos de carbono sin preocuparme	3	2	1	0	0	0
IC 2	Pienso que mis muslos son demasiado gruesos	0	0	0	1	2	3
O 2	Pienso en ponerme a dieta	0	0	0	1	2	3
IC 3	Creo que mi estómago tiene el tamaño adecuado	3	2	1	0	0	0
O3	Me siento muy culpable cuando como en exceso	0	0	0	1	2	3
IC 4	Me siento satisfecho con mi figura	3	2	1	0	0	0
O 4	Me aterroriza la idea de engordar	0	0	0	1	2	3
IC 5	Me gusta la forma de mi trasero	3	2	1	0	0	0
O 5	Exagero o doy demasiada importancia al peso	0	0	0	1	2	3
IC 6	Creo que mis caderas son demasiado anchas	0	0	0	1	2	3
O 6	Estoy preocupado porque querría ser una persona más delgada	0	0	0	1	2	3
IC 7	Creo que el tamaño de mis muslos es el adecuado	3	2	1	0	0	0
O 7	Si engordo un kilo, me preocupa que pueda seguir ganando peso	0	0	0	1	2	3
IC 8	Creo que mi trasero es demasiado grande	0	0	0	1	2	3
IC 9	Creo que mis caderas tienen el tamaño adecuado	3	2	1	0	0	0



XUNTA DE GALICIA  
CONSELLERÍA DE SANIDADE  
Secretaría Xeral

Comité Ético de Investigación Clínica de Galicia  
Edificio Administrativo de San Lázaro  
15781 SANTIAGO DE COMPOSTELA  
Tlf: 881 546425 Fax: 881 541804  
ceic@sergas.es

galicia

## DICTAMEN DEL COMITÉ ÉTICO DE INVESTIGACIÓN CLÍNICA DE GALICIA

Paula M. López Vázquez, Secretaria del Comité Ético de Investigación Clínica de Galicia

### CERTIFICA:

Que este Comité evaluó en su reunión del día 26/04/2012, la enmienda del estudio:

**Título:** Riesgo cardiovascular, comorbilidad, calidad de vida en personas de 40 y más años de edad en el municipio de Cambre

**Versión Enmienda:** modificación del protocolo y de los consentimientos de abril 2012

**Promotor:** Salvador Pita Fernández

**Código de Registro CEIC de Galicia:** 2008/264

Y que este Comité acepta de conformidad con sus procedimientos normalizados de trabajo y tomando en cuenta los requisitos éticos, metodológicos y legales exigibles a los estudios de investigación con seres humanos, sus muestras o registros, que dicha enmienda sea incorporada al estudio de investigación que se está realizando en los centros aprobados.

En Santiago de Compostela a 03 de mayo de 2012

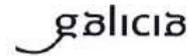
La Secretaria

Paula M. López Vázquez



XUNTA DE GALICIA  
CONSELLERÍA DE SANIDADE  
Secretaría Xeral

Comité Ético de Investigación Clínica de Galicia  
Edificio Administrativo de San Lázaro  
15781 SANTIAGO DE COMPOSTELA  
Tf: 881 546425 Fax: 881 541804  
ceic@sergas.es



## DICTAMEN DEL COMITÉ ÉTICO DE INVESTIGACIÓN CLÍNICA DE GALICIA

Juan Manuel Vázquez Lago, Secretario suplente del Comité Ético de Investigación Clínica de Galicia

### CERTIFICA:

Que este Comité evaluó en su reunión del día 20/12/2012 el estudio:

**Título:** Valoración del estado nutricional y autopercepción de la imagen corporal en una muestra aleatoria poblacional de 40 y más años en el municipio de Cambre

**Promotor:** Alba Arceo Vilas

**Tipo de estudio:** No-EPA

**Version:**

**Código do Promotor:**

**Código de Registro CEIC de Galicia:** 2012/390

Y que este Comité de conformidad con sus Procedimientos Normalizados de Trabajo y tomando en cuenta los requisitos éticos, metodológicos y legales exigibles a los estudios de investigación con seres humanos, sus muestras o registros, emite un **DICTAMEN FAVORABLE** al estudio propuesto y que se llevará a cabo en:

Centros	Investigadores principales
C.H. Universitario de A Coruña	Alba Arceo Vilas, Salvador Pita Fernández

En Santiago de Compostela a 07 de enero de 2013  
El Secretario suplente



Juan Manuel Vázquez Lago

## ÍNDICE DE PALABRAS SIGNIFICATIVAS

Aprendizaje automático, 22, 42, 43, 44, 45, 46, 65, 68, 113, 128, 134

Dieta mediterránea, 1, 3, 9, 11, 15, 16, 18, 19, 21, 22, 35, 38, 41, 47, 50, 61, 62, 64, 107, 108, 110, 111, 112, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 134, 137, 140, 150, 152, 153

Elastic Net, 44, 46, 89, 90, 125

Estado nutricional, 9, 11, 15, 18, 20, 26, 29, 31, 47, 50, 60, 61, 103, 104, 106, 107, 119, 134, 137, 139, 150, 152, 153

Inteligencia artificial, 9, 11, 13, 15, 16, 24, 42, 43, 44, 47, 137

kNN, 24, 46, 82

Machine learning, 9, 11, 13, 15, 16, 18, 24, 42, 66, 68, 70, 74, 112, 125, 131, 134

SVM, 18, 25, 44, 45, 76, 77, 78, 79, 80, 125, 126, 129, 131, 134, 143, 148

Obesidad, 20, 26, 27, 28, 31, 41, 57, 58, 59, 60, 97, 99, 138, 149

Random Forest, 18, 44, 45, 81, 125, 149

Sobrepeso, 26, 27, 28, 31, 57, 58, 97, 138, 149